THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MMAT5520

Differential Equations and Linear Algebra

Contents

1 Ordinary differential equations of first-order

An equation of the form

$$F(x, y, y', \cdots, y^{(n)}) = 0, \; x \in (a, b),$$

where $y = y(x), y' = \frac{dy}{dx}, \cdots, y^{(n)} = \frac{d^n y}{dx^n}$ is called an **ordinary differential equation (ODE)** of the function y.

Examples:

1. $y' - 4y = 0,$
2. $y'' - 3x^2 y' + 4\sqrt{x}y - 5e^{3x} + 1 = 0,$
3. $x^3 y'' + y \sin y' = 0.$

The **order** of the ODE is the order of the highest derivative in the equation. In solving ODE's, we are interested in the following problems:

- **Initial value problem(IVP)**: to find solutions $y(x)$ which satisfies given initial value conditions, e.g. $y(x_0) = y_0$, $y'(x_0) = y'_0$ for some constants y_0, y'_0.
- **Boundary value problem(BVP)**: to find solutions $y(x)$ which satisfies given boundary value conditions, e.g. $y(x_0) = y_0, y(x_1) = y_1$ for some constants y_0, y_1

An ODE is linear if it can be written as the form

$$p_n(x)y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(t)y = g(x), \; p_n(x) \neq 0.$$

The linear ODE is called **homogeneous** if $g(x) \equiv 0$, **nonhomogeneous**, otherwise. If an ODE is not of the above form, we call it a **non-linear** ODE.

1.1 First-order linear ODE

The general form of a **first-order linear ODE** is

$$y' + p(x)y = g(x).$$

The basic principle to solve a first-order linear ODE is to make left hand side a derivative of an expression by multiplying both sides by a suitable factor called an **integrating factor**. To find integrating factor, multiply both sides of the equation by $e^{f(x)}$, where $f(x)$ is to be determined, we have

$$e^{f(x)}y' + e^{f(x)}p(x)y = g(x)e^{f(x)}.$$

Now, if we choose $f(x)$ so that $f'(x) = p(x)$, then the left hand side becomes

$$e^{f(x)}y' + e^{f(x)}f'(x)y = \frac{d}{dx}\left(e^{f(x)}y\right).$$

Thus we may take

$$f(x) = \int p(x)dx$$

and the equation can be solved easily as follow.

$$\begin{aligned}
y' + p(x)y &= g(x) \\
e^{\int p(x)dx}\frac{dy}{dx} + e^{\int p(x)dx}p(x)y &= g(x)e^{\int p(x)dx} \\
\frac{d}{dx}e^{\int p(x)dx}y &= g(x)e^{\int p(x)dx} \\
e^{\int p(x)dx}y &= \int \left(g(x)e^{\int p(x)dx}\right) dx \\
y &= e^{-\int p(x)dx}\int \left(g(x)e^{\int p(x)dx}\right) dx
\end{aligned}$$

Note: Integrating factor is not unique. One may choose an arbitrary integration constant for $\int p(x)dx$. Any primitive function of $p(x)$ gives an integrating factor for the equation.

Example 1.1.1. *Find the general solution of* $y' + 2xy = 0$.

Solution: Multiplying both sides by e^{x^2}, we have

$$\begin{aligned}
e^{x^2}\frac{dy}{dx} + e^{x^2}2xy &= 0 \\
\frac{d}{dx}e^{x^2}y &= 0 \\
e^{x^2}y &= C \\
y &= Ce^{-x^2}
\end{aligned}$$

□

Example 1.1.2. *Solve* $(x^2-1)y' + xy = 2x$, $x > 1$.

Solution: Dividing both sides by $x^2 - 1$, the equation becomes

$$\frac{dy}{dx} + \frac{x}{x^2-1}y = \frac{2x}{x^2-1}.$$

Now

$$-\int \frac{x}{x^2-1}dx = \frac{1}{2}\ln(x^2-1) + C$$

Thus we multiply both sides of the equation by

$$\exp\left(\frac{1}{2}\ln(x^2-1)\right) = (x^2-1)^{\frac{1}{2}}$$

and get

$$\begin{aligned}
(x^2-1)^{\frac{1}{2}}\frac{dy}{dx} + \frac{x}{(x^2-1)^{\frac{1}{2}}}y &= \frac{2x}{(x^2-1)^{\frac{1}{2}}} \\
\frac{d}{dx}\left((x^2-1)^{\frac{1}{2}}y\right) &= \frac{2x}{(x^2-1)^{\frac{1}{2}}} \\
(x^2-1)^{\frac{1}{2}}y &= \int \frac{2x}{(x^2-1)^{\frac{1}{2}}}dx \\
y &= (x^2-1)^{-\frac{1}{2}}\left(2(x^2-1)^{\frac{1}{2}} + C\right) \\
y &= 2 + C(x^2-1)^{-\frac{1}{2}}
\end{aligned}$$

$\square$

Example 1.1.3. *Solve* $y' - y\tan x = 4\sin x$, $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

Solution: An integrating factor is

$$\exp(-\int \tan x dx) = \exp(\ln(\cos x)) = \cos x.$$

Multiplying both sides by $\cos x$, we have

$$\begin{aligned}
\cos x\frac{dy}{dx} - y\sin x &= 4\sin x\cos x \\
\frac{d}{dx}(y\cos x) &= 2\sin 2x \\
y\cos x &= \int 2\sin 2x dx \\
y\cos x &= -\cos 2x + C \\
y &= \frac{C - \cos 2x}{\cos x}
\end{aligned}$$

$\square$

Example 1.1.4. *A tank contains 1L of a solution consisting of 100 g of salt dissolved in water. A salt solution of concentration of 20 gL*$^{-1}$ *is pumped into the tank at the rate of 0.02 Ls*$^{-1}$*, and the mixture, kept uniform by stirring, is pumped out at the same rate. How long will it be until only 60 g of salt remains in the tank?*

Solution: Suppose there is x g of salt in the solution at time t s. Then x follows the following differential equation

$$\frac{dx}{dt} = 0.02(20 - x)$$

Multiplying the equation by $e^{0.02t}$, we have

$$\begin{aligned}
\frac{dx}{dt} + 0.02x &= 0.4 \\
e^{0.02t}\frac{dx}{dt} + 0.02e^{0.02t}x &= 0.4e^{0.02t} \\
\frac{d}{dt}e^{0.02t}x &= \int 0.4e^{0.02t}dt \\
e^{0.02t}x &= 20e^{0.02t} + C \\
x &= 20 + Ce^{-0.02t}
\end{aligned}$$

Since $x(0) = 100$, $C = 80$. Thus the time taken in second until 60 g of salt remains in the tank is

$$\begin{aligned}
60 &= 20 + 80e^{-0.02t} \\
e^{0.02t} &= 2 \\
t &= 50\ln 2
\end{aligned}$$

$\square$

Example 1.1.5. *David would like to buy a home. He had examined his budget and determined that he can afford monthly payments of \$20,000. If the annual interest is 6% compounded continuously, and the term of the loan is 20 years, what amount can he afford to borrow?*

Solution: Let \$$y$ be the remaining loan amount after t months. Then

$$\begin{aligned}
\frac{dy}{dt} &= \frac{0.06}{12}y - 20,000 \\
\frac{dy}{dt} - 0.005y &= -20,000 \\
e^{-0.005t}\frac{dy}{dt} - 0.005ye^{-0.005t} &= -20,000e^{-0.005t} \\
\frac{d}{dt}(e^{-0.005t}y) &= -20,000e^{-0.005t} \\
e^{-0.005t}y &= \frac{-20,000e^{-0.005t}}{-0.005} + C \\
y &= 4,000,000 + Ce^{0.005t}
\end{aligned}$$

Since the term of the loan is 20 years, $y(240) = 0$ and thus

$$\begin{aligned}
4,000,000 + Ce^{0.005\times 240} &= 0 \\
C &= -\frac{4,000,000}{e^{1.2}} \\
&= -1,204,776.85
\end{aligned}$$

Therefore the amount that David can afford to borrow is

$$\begin{aligned}
y(0) &= 4,000,000 - 1,204,776.85e^{0.005(0)} \\
&= 2,795,223.15
\end{aligned}$$

Note: The total amount that David pays is \$240 × 20,000 = \$4,800,000. □

Exercise 1.1

1. Find the general solutions of the following first order linear differential equations.

   (a) $y' + y = 4e^{3x}$
   (b) $3xy' + y = 12x$
   (c) $y' + 3x^2y = x^2$
   (d) $x^2y' + xy = 1$
   (e) $xy' + y = \sqrt{x}$
   (f) $xy' = y + x^2\sin x$
   (g) $(x+1)y' - 2y = (x+1)^{\frac{7}{2}}$
   (h) $y'\cos x + y\sin x = 1$
   (i) $xy' + (3x+1)y = e^{-3x}$

2. Solve the following initial value problems.

   (a) $y' - y = e^{2x}$; $y(0) = 1$
   (b) $y' = (1-y)\cos x$; $y(\pi) = 2$
   (c) $(x^2+4)y' + 3xy = 3x$; $y(0) = 3$
   (d) $(x+1)y' + y = \ln x$; $y(1) = 10$
   (e) $x^2y' + 2xy = \ln x$; $y(1) = 2$
   (f) $xy' + y = \sin x$; $y(\pi) = 1$

1.2 Separable equations

A **separable equation** is an equation of the form

$$\frac{dy}{dx} = f(x)g(y).$$

It can be solved as follows

$$\begin{aligned}\frac{dy}{g(y)} &= f(x)dx \\ \int \frac{dy}{g(y)} &= \int f(x)dx\end{aligned}$$

Example 1.2.1. *Find the general solution of* $y' = 3x^2y$.

Solution:

$$\begin{aligned}\frac{dy}{y} &= 3x^2dx \\ \int \frac{dy}{y} &= \int 3x^2dx \\ \ln y &= x^3 + C' \\ y &= Ce^{x^3} \quad \text{where } C = e^{C'}\end{aligned}$$

□

Example 1.2.2. *Solve* $2\sqrt{x}\dfrac{dy}{dx} = y^2 + 1,\ x > 0$.

Solution:

$$\begin{aligned}\frac{dy}{y^2+1} &= \frac{dx}{2\sqrt{x}} \\ \int \frac{dy}{y^2+1} &= \int \frac{dx}{2\sqrt{x}} \\ \tan^{-1} y &= \sqrt{x} + C \\ y &= \tan(\sqrt{x} + C)\end{aligned}$$

□

Example 1.2.3. *Solve the initial value problem* $\dfrac{dy}{dx} = \dfrac{x}{y + x^2y},\ y(0) = -1$.

Solution:

$$\begin{aligned}\frac{dy}{dx} &= \frac{x}{y(1+x^2)} \\ \int ydy &= \int \frac{x}{1+x^2}dx \\ \frac{y^2}{2} &= \frac{1}{2}\int \frac{1}{1+x^2}d(1+x^2) \\ y^2 &= \ln(1+x^2) + C\end{aligned}$$

Since $y(0) = -1$, $C = 1$. Thus

$$\begin{aligned}y^2 &= 1 + \ln(1+x^2) \\ y &= -\sqrt{1 + \ln(1+x^2)}\end{aligned}$$

□

Example 1.2.4. *(Logistic equation) Solve the initial value problem for the* **logistic equation**

$$\frac{dy}{dt} = ry(1 - y/K),\ y(0) = y_0$$

where r and K are constants.

Solution:

$$\begin{aligned} \frac{dy}{y(1-y/K)} &= rdt \\ \int \frac{dy}{y(1-y/K)}dt &= \int rdt \\ \int \left(\frac{1}{y} + \frac{1/K}{1-y/K}\right) dt &= rt \\ \ln y - \ln(1-y/K) &= rt + C \\ \frac{y}{1-y/K} &= e^{rt+C} \\ y &= \frac{Ke^{rt+C}}{K+e^{rt+C}} \end{aligned}$$

To satisfy the initial condition, we set

$$e^C = \frac{y_0}{1 - y_0/K}$$

and obtain

$$y = \frac{y_0 K}{y_0 + (K - y_0)e^{-rt}}.$$

Note: When $t \to \infty$,

$$\lim_{t\to\infty} y(t) = K.$$

□

Exercise 1.2

1. Find the general solution of the following separable equations.

   (a) $y' + 2xy^2 = 0$
   (b) $y' = 3\sqrt{xy}$
   (c) $y' = 6x(y-1)^{\frac{2}{3}}$
   (d) $y' = y\sin x$
   (e) $yy' = x(y^2+1)$
   (f) $y' = 1 + x + y + xy$

2. Solve the following initial value problems.

   (a) $xy' - y = 2x^2y$; $y(1) = 1$
   (b) $y' = ye^x$; $y(0) = 2e$
   (c) $2yy' = \dfrac{x}{\sqrt{x^2-16}}$; $y(5) = 2$
   (d) $y' = 4x^3y - y$; $y(1) = -3$
   (e) $y'\tan x = y$; $y(\frac{\pi}{2}) = \frac{\pi}{2}$
   (f) $y' = 3x^2(y^2+1)$; $y(0) = 1$

3. Solve the logistic equation

   $$\frac{dy}{dx} = 0.08y\left(1 - \frac{y}{1000}\right)$$

   with initial condition $y(0) = 100$.

1.3 Exact equations

We say that the equation

$$M(x,y)dx + N(x,y)dy = 0 \tag{1.4.1}$$

is **exact** if

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

In this case, there exists a function $f(x,y)$ such that

$$\begin{cases} \frac{\partial f}{\partial x} = M \\ \frac{\partial f}{\partial y} = N \end{cases}$$

Then the differential equation can be written as

$$\begin{aligned} \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy &= 0 \\ df(x,y) &= 0 \end{aligned}$$

Therefore the general solution of the differential equation is

$$f(x,y) = C.$$

To find $f(x,y)$, first note that

$$\frac{\partial f}{\partial x} = M.$$

Hence

$$f(x,y) = \int M(x,y)dx + g(y).$$

Differentiating both sides with respect to y, we have

$$N(x,y) = \frac{\partial}{\partial y}\int M(x,y)dx + g'(y)$$

since

$$\frac{\partial f}{\partial y} = N.$$

Now

$$N(x,y) - \frac{\partial}{\partial y}\int M(x,y)dx$$

is independent of x (why?). Therefore

$$g(y) = \int \left(N(x,y) - \frac{\partial}{\partial y}\int M(x,y)dx\right)dy$$

and we obtain

$$\begin{aligned} f(x,y) &= \int M(x,y)dx + g(y) \\ &= \int M(x,y)dx + \int \left(N(x,y) - \frac{\partial}{\partial y}\int M(x,y)dx\right)dy \end{aligned}$$

Remark: Equation (1.4.1) is exact when $\mathbf{F} = (M(x,y), N(x,y))$ defines a conservative vector field. The function $f(x,y)$ is called a potential function for $\mathbf{F}$.

Example 1.3.1. *Solve* $(4x+y)dx+(x-2y)dy=0$.

Solution: Since

$$\frac{\partial}{\partial y}(4x+y)=1=\frac{\partial}{\partial x}(x-2y),$$

the equation is exact. We need to find $F(x,y)$ such that

$$\frac{\partial F}{\partial x}=M \text{ and } \frac{\partial F}{\partial y}=N.$$

Now

$$\begin{aligned} F(x,y) &= \int(4x+y)dx \\ &= 2x^2+xy+g(y) \end{aligned}$$

To determine $g(y)$, what we want is

$$\begin{aligned} \frac{\partial F}{\partial y} &= x-2y \\ x+g'(y) &= x-2y \\ g'(y) &= -2y \end{aligned}$$

Therefore we may choose $g(y)=-y^2$ and the solution is

$$F(x,y)=2x^2+xy-y^2=C.$$

□

Example 1.3.2. *Solve* $\dfrac{dy}{dx}=\dfrac{e^y+x}{e^{2y}-xe^y}$.

Solution: Rewrite the equation as

$$(e^y+x)dx+(xe^y-e^{2y})dy=0.$$

Since

$$\frac{\partial}{\partial y}(e^y+x)=e^y=\frac{\partial}{\partial x}(xe^y-e^{2y}),$$

the equation is exact. Set

$$\begin{aligned} F(x,y) &= \int(e^y+x)dx \\ &= xe^y+\frac{1}{2}x^2+g(y) \end{aligned}$$

We want

$$\begin{aligned} \frac{\partial F}{\partial y} &= xe^y-e^{2y} \\ xe^y+g'(y) &= xe^y-e^{2y} \\ g'(y) &= -e^{2y} \end{aligned}$$

Therefore we may choose $g(y) = -\frac{1}{2}e^{2y}$ and the solution is

$$xe^y + \frac{1}{2}x^2 - \frac{1}{2}e^{2y} = C.$$

□

When the equation is not exact, sometimes it is possible to convert it to an exact equation by multiplying it by a suitable **integrating factor**. Unfortunately, there is no systematic way of finding integrating factor in general.

Example 1.3.3. *Show that* $\mu(x, y) = x$ *is an integrating factor of* $(3xy+y^2)dx+(x^2+xy)dy = 0$ *and then solve the equation.*

Solution: Multiplying the equation by x reads

$$(3x^2y + xy^2)dx + (x^3 + x^2y)dy = 0.$$

Now

$$\begin{aligned} \frac{\partial}{\partial y}(3x^2y + xy^2) &= 3x^2 + 2xy \\ \frac{\partial}{\partial x}(x^3 + x^2y) &= 3x^2 + 2xy \end{aligned}$$

Thus the above equation is exact and x is an integrating factor. To solve the equation, set

$$\begin{aligned} F(x, y) &= \int (3x^2y + xy^2)dx \\ &= x^3y + \frac{1}{2}x^2y^2 + g(y) \end{aligned}$$

Now we want

$$\begin{aligned} \frac{\partial F}{\partial y} &= x^3 + x^2y \\ x^3 + x^2y + g'(y) &= x^3 + x^2y \\ g'(y) &= 0 \end{aligned}$$

Therefore $g(y)$ is constant and the solution is

$$x^3y + \frac{1}{2}x^2y^2 = C.$$

□

Note: The equation in Example 1.3.3 is also a homogenous equation which will be discussed in Section 1.4.

Example 1.3.4. *Show that* $\mu(x, y) = y$ *is an integrating factor of* $ydx + (2x - e^y)dy = 0$ *and then solve the equation.*

Solution: Multiplying the equation by y reads

$$y^2dx + (2xy - ye^y)dy = 0.$$

Now

$$\begin{aligned}\frac{\partial}{\partial y}y^2 &= 2y\\ \frac{\partial}{\partial x}(2xy - ye^y) &= 2y\end{aligned}$$

Thus the above equation is exact and y is an integrating factor. To solve the equation, set

$$\begin{aligned}F(x,y) &= \int y^2 dx\\ &= xy^2 + g(y)\end{aligned}$$

Now we want

$$\begin{aligned}\frac{\partial F}{\partial y} &= 2xy - ye^y\\ 2xy + g'(y) &= 2xy - ye^y\\ g'(y) &= -ye^y\\ g(y) &= -\int ye^y dy\\ &= -\int yde^y\\ &= -ye^y + \int e^y dy\\ &= -ye^y + e^y + C'\end{aligned}$$

Therefore the solution is

$$xy^2 - ye^y + e^y = C.$$

□

Exercise 1.3

1. For each of the following equations, show that it is exact and find the general solution.
   - (a) $(5x+4y)dx + (4x-8y^3)dy = 0$
   - (b) $(3x^2+2y^2)dx + (4xy+6y^2)dy = 0$
   - (c) $(3xy^2 - y^3)dx + (3x^2y - 3xy^2)dy = 0$
   - (d) $(1 + ye^{xy})dx + (2y + xe^{xy})dy = 0$
   - (e) $(x^3 + \frac{y}{x})dx + (y^2 + \ln x)dy = 0$
   - (f) $(\cos x + \ln y)dx + (\frac{x}{y} + e^y)dy = 0$
2. For each of the following differential equations, find the value of k so that the equation is exact and solve the equation
   - (a) $(2xy^2 - 3)dx + (kx^2y + 4)dy = 0$
   - (b) $(6xy - y^3)dx + (4y + 3x^2 + kxy^2)dy = 0$
   - (c) $(2xy^2 + 3x^2)dx + (2x^ky + 4y^3)dy = 0$
   - (d) $(3x^2y^3 + y^k)dx + (3x^3y^2 + 4xy^3)dy = 0$
3. For each of the following differential equations, show that the given function μ is an integrator of the equation and then solve the equation.
   - (a) $(3xy + y^2)dx + (x^2 + xy)dy = 0$; $\mu(x) = x$
   - (b) $ydx - xdy = 0$; $\mu(y) = \frac{1}{y^2}$
   - (c) $ydx + x(1 + y^3)dy = 0$; $\mu(x,y) = \dfrac{1}{xy}$
   - (d) $(x - y)dx + (x + y)dy = 0$; $\mu(x,y) = \dfrac{1}{x^2 + y^2}$

1.4 Homogeneous equations

A first order equation is **homogeneous** if it can be written as

$$\frac{dy}{dx} = f\left(\frac{y}{x}\right).$$

The above equation can be solved by the substitution $u = y/x$. Then $y = xu$ and

$$\frac{dy}{dx} = u + x\frac{du}{dx}.$$

Therefore the equation reads

$$\begin{aligned} u + x\frac{du}{dx} &= f(u) \\ \frac{du}{f(u) - u} &= \frac{dx}{x} \end{aligned}$$

which becomes a separable equation.

Example 1.4.1. *Solve* $\frac{dy}{dx} = \frac{x^2 + y^2}{2xy}$.

Solution: Rewrite the equation as

$$\frac{dy}{dx} = \frac{1 + (y/x)^2}{2y/x}$$

which is a homogeneous equation. Using substitution $y = xu$, we have

$$\begin{aligned} u + x\frac{du}{dx} &= \frac{dy}{dx} = \frac{1 + u^2}{2u} \\ x\frac{du}{dx} &= \frac{1 + u^2 - 2u^2}{2u} \\ \frac{2udu}{1 - u^2} &= \frac{dx}{x} \\ \int \frac{2udu}{1 - u^2} &= \int \frac{dx}{x} \\ -\ln(1 - u^2) &= \ln x + C' \\ (1 - u^2)x &= e^{-C'} \\ x^2 - y^2 - Cx &= 0 \end{aligned}$$

where $C = e^{-C'}$. □

Example 1.4.2. *Solve* $(y + 2xe^{-y/x})dx - xdy = 0$.

Solution: Rewrite the equation as

$$\frac{dy}{dx} = \frac{y + 2xe^{-y/x}}{x} = \frac{y}{x} + 2e^{-y/x}.$$

Let $u = y/x$, we have

$$\begin{aligned} u + x\frac{du}{dx} &= \frac{dy}{dx} = u + 2e^{-u} \\ x\frac{du}{dx} &= 2e^{-u} \\ e^u du &= 2\frac{dx}{x} \\ \int e^u du &= \int 2\frac{dx}{x} \\ e^u &= 2\ln x + C \\ e^{y/x} - 2\ln x &= C \end{aligned}$$

☐

Exercise 1.4

1. Find the general solution of the following homogeneous equations.

(a) $y' = \dfrac{x^2 + 2y^2}{2xy}$

(b) $xy' = y + \sqrt{x^2 - y^2}$

(c) $xy' = y + 2\sqrt{xy}$

(d) $x(x+y)y' = y(x-y)$

(e) $x^2y' = xy + y^2$

(f) $x^2y' = xy + x^2e^{\frac{y}{x}}$

1.5 Bernoulli's equations

An equation of the form

$$y' + p(x)y = q(x)y^n, \quad n \neq 0, 1,$$

is called a **Bernoulli's equation**. It is a non-linear equation and $y(x) = 0$ is always a solution when $n > 0$. To find a non-trivial solution, we use the substitution

$$u = y^{1-n}.$$

Then

$$\begin{aligned} \frac{du}{dx} &= (1-n)y^{-n}\frac{dy}{dx} \\ &= (1-n)y^{-n}\left(-p(x)y + q(x)y^n\right) \\ \frac{du}{dx} + (1-n)p(x)y^{1-n} &= (1-n)q(x) \\ \frac{du}{dx} + (1-n)p(x)u &= (1-n)q(x) \end{aligned}$$

which is a linear differential equation of u.
Note: Don't forget that $y(x) = 0$ is always a solution to the Bernoulli's equation when $n > 0$.

Example 1.5.1. *Solve* $\dfrac{dy}{dx} - y = e^{-x}y^2$.

Solution: Let $u = y^{1-2} = y^{-1}$,

$$\begin{aligned} \frac{du}{dx} &= -y^{-2}\frac{dy}{dx} \\ &= -y^{-2}\left(y + e^{-x}y^2\right) \\ \frac{du}{dx} + y^{-1} &= -e^{-x} \\ \frac{du}{dx} + u &= -e^{-x} \end{aligned}$$

which is a linear equation of u. Multiplying both side by e^x, we have

$$\begin{aligned} e^x\frac{du}{dx}+e^xu &= -1 \\ \frac{d}{dx}(e^xu) &= -1 \\ e^xu &= -x+C \\ u &= (C-x)e^{-x} \\ y^{-1} &= (C-x)e^{-x} \end{aligned}$$

Therefore the general solution is

$$y=\frac{e^x}{C-x}\text{ or }y=0$$

□

Example 1.5.2. *Solve* $x\dfrac{dy}{dx}+y=xy^3$.

Solution: Let $u=y^{1-3}=y^{-2}$,

$$\begin{aligned} \frac{du}{dx} &= -2y^{-3}\frac{dy}{dx} \\ \frac{du}{dx} &= -\frac{2y^{-3}}{x}\left(-y+xy^3\right) \\ \frac{du}{dx}-\frac{2y^{-2}}{x} &= -2 \\ \frac{du}{dx}-\frac{2u}{x} &= -2 \end{aligned}$$

which is a linear equation of u. To solve it, multiply both side by $\exp(-\int 2x^{-1}dx)=x^{-2}$, we have

$$\begin{aligned} x^{-2}\frac{du}{dx}-2x^{-3}u &= -2x^{-2} \\ \frac{d}{dx}(x^{-2}u) &= -2x^{-2} \\ x^{-2}u &= 2x^{-1}+C \\ u &= 2x+Cx^2 \\ y^{-2} &= 2x+Cx^2 \\ y^2 &= \frac{1}{2x+Cx^2}\text{ or }y=0 \end{aligned}$$

□

Exercise 1.5

1. Find the general solution of the following Bernoulli's equations.

   (a) $xy'+y=x^2y^2$ (b) $x^2y'+2xy=5y^4$ (c) $xy'=y(x^2y-1)$

1.6 Substitution

In this section, we give some examples of differential equations that can be transformed to one of the forms in the previous sections by a suitable substitution.

Example 1.6.1. *Use the substitution* $u = \ln y$ *to solve* $xy' - 4x^2y + 2y\ln y = 0$.

Solution: Substitute

$$u' = \frac{y'}{y}$$

into the equation, we have

$$\begin{aligned} xy' - 4x^2y + 2y\ln y &= 0 \\ x\left(\frac{y'}{y}\right) + 2\ln y &= 4x^2 \\ xu' + 2u &= 4x^2 \\ x^2u' + 2xu &= 4x^3 \\ \frac{d}{dx}x^2u &= 4x^3 \\ x^2u &= \int 4x^3dx \\ x^2u &= x^4 + C \\ u &= x^2 + \frac{C}{x^2} \\ y &= \exp\left(x^2 + \frac{C}{x^2}\right) \end{aligned}$$

□

Example 1.6.2. *Use the substitution* $u = e^{2y}$ *to solve* $2xe^{2y}y' = 3x^4 + e^{2y}$.

Solution: Substitute

$$u' = 2e^{2y}y'$$

into the equation, we have

$$\begin{aligned} 2xe^{2y}y' &= 3x^4 + e^{2y} \\ xu' - u &= 3x^4 \\ \frac{u'}{x} - \frac{u}{x^2} &= 3x^2 \\ \frac{d}{dx}\left(\frac{u}{x}\right) &= 3x^2 \\ \frac{u}{x} &= \int 3x^2dx \\ \frac{u}{x} &= x^3 + C \\ e^{2y} &= x^4 + Cx \\ y &= \frac{1}{2}\ln|x^4 + Cx| \end{aligned}$$

$\square$

An equation of the form

$$y' + p_1(x)y + p_2(x)y^2 = q(x)$$

is called a **Riccati's equation**. If we know that $y(x) = y_p(x)$ is a particular solution, then the equation can be transformed, using the substitution

$$y = y_p + \frac{1}{u}$$

to a linear equation of u.

Example 1.6.3. *Solve the Riccati's equation* $y' - \dfrac{y}{x} = 1 - \dfrac{y^2}{x^2}$ *given that* $y = x$ *is a particular solution.*

Solution: Let

$$y = x + \frac{1}{u}.$$

We have

$$\begin{array}{rcl}
\dfrac{dy}{dx} & = & 1 - \dfrac{1}{u^2}\dfrac{du}{dx} \\
\dfrac{1}{x}y + 1 - \dfrac{1}{x^2}y^2 & = & 1 - \dfrac{1}{u^2}\dfrac{du}{dx} \\
\dfrac{1}{u^2}\dfrac{du}{dx} & = & \dfrac{1}{x^2}\left(x + \dfrac{1}{u}\right)^2 - \dfrac{1}{x}\left(x + \dfrac{1}{u}\right) \\
\dfrac{1}{u^2}\dfrac{du}{dx} & = & \dfrac{1}{xu} + \dfrac{1}{x^2u^2} \\
\dfrac{du}{dx} - \dfrac{1}{x}u & = & \dfrac{1}{x^2}
\end{array}$$

which is a linear equation of u. An integrating factor is

$$\exp\left(-\int \frac{1}{x}dx\right) = \exp(-\ln x) = x^{-1}.$$

Thus

$$\begin{array}{rcl}
x^{-1}\dfrac{du}{dx} - x^{-2}u & = & x^{-3} \\
\dfrac{d}{dx}(x^{-1}u) & = & x^{-3} \\
x^{-1}u & = & -\dfrac{1}{2x^2} + C' \\
u & = & -\dfrac{1}{2x} + C'x \\
u & = & \dfrac{Cx^2 - 1}{2x}
\end{array}$$

Therefore the general solution is

$$y = x + \frac{2x}{Cx^2 - 1} \text{ or } y = x.$$

$\square$

Example 1.6.4. *Solve the Riccati's equation $y' = 1 + x^2 - 2xy + y^2$ given that $y = x$ is a particular solution.*

Solution: Using the substitution

$$y = x + \frac{1}{u}.$$

We have

$$\begin{aligned} \frac{dy}{dx} &= 1 - \frac{1}{u^2}\frac{du}{dx} \\ 1 + x^2 - 2xy + y^2 &= 1 - \frac{1}{u^2}\frac{du}{dx} \\ 1 + x^2 - 2x(x + \frac{1}{u}) + (x + \frac{1}{u})^2 &= 1 - \frac{1}{u^2}\frac{du}{dx} \\ \frac{du}{dx} &= -1 \\ u &= C - x \end{aligned}$$

Therefore the general solution is

$$y = x + \frac{1}{C - x} \text{ or } y = x.$$

□

Exercise 1.6

1. Solve the following differential equations by using the given substitution.
   (a) $xy' - 4x^2y + 2y\ln y = 0$; $u = \ln y$
   (b) $y' = \sqrt{x + y}$; $u = x + y$
   (c) $y' = (x + y + 3)^2$; $u = x + y + 3$
   (d) $y' + e^y + 1 = 0$; $u = e^{-y}$

2. Solve the following Riccati's equations by the substitution $y = y_1 + \frac{1}{u}$ with the given particular solution $y_1(x)$.
   (a) $x^3y' = y^2 + x^2y - x^2$; $y_1(x) = x$
   (b) $x^2y' - x^2y^2 = -2$; $y_1(x) = \frac{1}{x}$

1.7 Reducible second-order equations

Some second-order differential equations can be reduced to an equation of first-order by a suitable substitution. First we consider the simplest case when the zeroth order derivative term y is missing.

Dependent variable y missing:

$$F(x, y', y'') = 0$$

The substitution

$$p = y', \; y'' = \frac{dp}{dx} = p',$$

reduces the equation into a first-order differential equation

$$F(x, p, p') = 0.$$

Example 1.7.1. *Solve* $xy'' + 2y' = 6x$*.*

Solution: Let $p = y'$. Then $y'' = p'$ and the equation reads

$$\begin{aligned} xp' + 2p &= 6x \\ x^2p' + 2xp &= 6x^2 \\ \frac{d}{dx}(x^2p) &= 6x^2 \\ x^2p &= 2x^3 + C_1 \\ y' &= 2x + C_1x^{-2} \\ y &= x^2 - C_1x^{-1} + C_2 \end{aligned}$$

□

Example 1.7.2 (Free falling with air resistance)**.** *The motion of an free falling object near the surface of the earth with air resistance proportional to velocity can be modeled by the equation*

$$y'' + \rho y' + g = 0,$$

where ρ *is a positive constant and* g *is the gravitational acceleration. Solve the equation with initial displacement* $y(0) = y_0$ *and initial velocity* $\frac{dy}{dt}(0) = v_0$*.*

Solution: Let $v = \frac{dy}{dt}$. Then $\frac{d^2y}{dt^2} = \frac{dv}{dt}$ and

$$\begin{aligned} \frac{dv}{dt} + \rho v + g &= 0 \\ \int_{v_0}^{v} \frac{dv}{\rho v + g} &= -\int_0^t dt \\ \frac{1}{\rho}\ln(\rho v + g) &= -t \\ \ln(\rho v + g) - \ln(\rho v_0 + g) &= -\rho t \\ \rho v + g &= (\rho v_0 + g)e^{-\rho t} \\ v &= (v_0 - v_\tau)e^{-\rho t} + v_\tau \end{aligned}$$

where

$$v_\tau = -\frac{g}{\rho} = \lim_{t\to\infty} v(t)$$

is the terminal velocity. Thus

$$\begin{aligned} \frac{dy}{dt} &= (v_0 - v_\tau)e^{-\rho t} + v_\tau \\ \int_{y_0}^{y} dy &= \int_0^t \left((v_0 - v_\tau)e^{-\rho t} + v_\tau\right) dt \\ y - y_0 &= \left[-\frac{1}{\rho}(v_0 - v_\tau)e^{-\rho t} + v_\tau t\right]_0^t \\ y &= \frac{1}{\rho}(v_0 - v_\tau)(1 - e^{-\rho t}) + v_\tau t + y_0 \end{aligned}$$

□

Independent variable x missing:
The substitution

$$p = y', \; y'' = \frac{dp}{dx} = \frac{dp}{dy}\frac{dy}{dx} = \frac{dp}{dy}p$$

reduces the equation to the first-order equation

$$F\left(y, p, p\frac{dp}{dy}\right) = 0.$$

Example 1.7.3. *Solve $yy'' = y'^2$.*

Solution: Let $p = y'$. Then $y'' = p\dfrac{dp}{dy}$ and the equation reads

$$\begin{aligned}
yp\frac{dp}{dy} &= p^2 \\
\frac{dp}{p} &= \frac{dy}{y} \\
\int \frac{dp}{p} &= \int \frac{dy}{y} \\
\ln p &= \ln y + C' \\
p &= C_1 y \\
\frac{dy}{dx} &= C_1 y \\
\int \frac{dy}{y} &= \int C_1 dx \\
\ln y &= C_1 x + C' \\
y &= C_2 e^{C_1 x}
\end{aligned}$$

□

Example 1.7.4 (Escape velocity)**.** *The motion of an object projected vertically without propulsion from the earth surface is modeled by*

$$\frac{d^2r}{dt^2} = -\frac{GM}{r^2},$$

where r is the distance from the earth's center, $G \approx 6.6726 \times 10^{-11} Nm^2kg^2$ is the constant of universal gravitation and $M \approx 5.975 \times 10^24 kg$ is the mass of the earth. Find the minimum initial velocity v_0 for the projectile to escape from the earth's gravity.

Solution: Let $v = \dfrac{dr}{dt}$. Then $\dfrac{d^2r}{dt^2} = \dfrac{dv}{dt} = \dfrac{dv}{dr}\dfrac{dr}{dt} = \dfrac{dv}{dr}v$ and

$$\begin{aligned}
v\frac{dv}{dr} &= -\frac{GM}{r^2} \\
\int_{v_0}^{v} v dv &= -\int_{r_0}^{r} \frac{GM}{r^2} dr \\
\frac{1}{2}(v^2 - v_0^2) &= GM\left(\frac{1}{r} - \frac{1}{r_0}\right) \\
v_0^2 &= v^2 + 2GM\left(\frac{1}{r_0} - \frac{1}{r}\right)
\end{aligned}$$

where $r_0 \approx 6.378 \times 10^6$m is the radius of the earth. In order to escape from earth's gravity, r can be arbitrary large and thus we must have

$$\begin{aligned} v_0^2 &\geq \frac{2GM}{r_0} + v^2 \geq \frac{2GM}{r_0} \\ v_0 &\geq \sqrt{\frac{2GM}{r_0}} \approx 11,180(ms^{-1}) \end{aligned}$$

□

Exercise 1.7

1. Find the general solution of the following differential equations by reducing them to first order equations.
   - (a) $yy'' + (y')^2 = 0$
   - (b) $y'' + 4y = 0$
   - (c) $xy'' + y' = 4x$
   - (d) $x^2y'' + 3xy' = 2$
   - (e) $yy'' + (y')^2 = yy'$
   - (f) $y'' = 2y(y')^3$
2. Find the general solution of the following differential equations.
   - (a) $y' = xy^3$
   - (b) $y' = \dfrac{x^2 + 2y}{x}$
   - (c) $y' = \dfrac{1 - 9x^2 - y}{x - 4y}$
   - (d) $xy' + 2y = 6x^2\sqrt{y}$
   - (e) $x^2y' - xy - y^2 = 0$
   - (f) $x^2y' + 2xy^2 = y^2$
   - (g) $y' = 1 + x^2 + y^2 + x^2y^2$
   - (h) $x^2y' + 2xy = x - 1$
   - (i) $(1 - x)y' + y - x = 0$
   - (j) $y' + \dfrac{6xy^3 + 2y^4}{9x^2y^2 + 8xy^3} = 0$
   - (k) $x^3y' = x^2y - y^3$
   - (l) $3xy' + x^3y^4 + 3y = 0$

2 Linear systems and matrices

The study of linear algebra is motivated by trying to understand the solutions to a **linear system** which is a system of m linear equations in n unknowns

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & = & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}\right. .$$

The above system of linear equations can be expressed in the following matrix form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

For each linear system, there associates an $m \times (n+1)$ matrix

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

which is called the **augmented matrix** of the system. The **solution set** of the system is the set

$$\{\mathbf{v} \in \mathbb{R}^n : \mathbf{A}\mathbf{v} = \mathbf{0}\}$$

of vectors in $\mathbb{R}^n$ which satisfy the system. Two linear systems are **equivalent** if they have the same solution set.

2.1 Gaussian elimination

The idea of solving a general linear system is to transform the system to an equivalent system of certain type whose solution set can be written down easily. The type of systems that we would transform a system to is called row echelon form. The system is transformed using elementary row operations. The process of obtaining row echelon form using elementary row operations is called **Gaussian elimination**. First we give the definition of row echelon form.

Definition 2.1.1 (Row echelon form)**.** *A matrix* $\mathbf{R}$ *is said to be in* **row echelon form** *if it satisfies the following three properties:*

1. *Every row of* $\mathbf{R}$ *that consists entirely of zeros lies beneath every row that contains a nonzero entry.*
2. *The first nonzero entry of each row of* $\mathbf{R}$ *is 1.*
3. *In each row of* $\mathbf{R}$ *that contains a nonzero entry, the number of leading zeros is strictly less than the number of leading zeros in the preceding rows.*

Example 2.1.2. *The following matrices are not in row echelon form.*

$\begin{pmatrix} 1 & 0 & 5 & -1 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$	*There is a nonzero row, the 3rd row, lies beneath a zero row, the 2nd row.*
$\begin{pmatrix} 1 & 3 & 0 & -2 & 4 \\ 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	*The first nonzero entry of the second row is not one.*
$\begin{pmatrix} 1 & 4 & -2 & 0 & 5 \\ 0 & 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & 0 & 4 \end{pmatrix}$	*The number of leading zeros in the 2nd row is not strictly less then the number of zeros in the 3rd row.*

Example 2.1.3. *The following matrices are in row echelon form.*

$$\begin{pmatrix} 1 & -1 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 4 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & -3 & 2 & -1 & 2 \\ 0 & 1 & 4 & 5 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

To write down the solution set of a linear system associated with a row echelon, we need to distinguish between leading variables and free variables.

Definition 2.1.4 (Leading and free variables)**.** *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix and* $\mathbf{b}$ *be an* m *column vector. Suppose the augmented matrix* $R = [\mathbf{A}|\mathbf{b}]$ *associated with the linear system* $\mathbf{Ax} = \mathbf{b}$*, where* $\mathbf{x} = (x_1, \ldots, x_n)^T$*, is in row echelon form. Then for* $1 \leq k \leq n$*,*

1. *if the k-th column contains a leading nonzero entries, then* x_k *is a* **leading variable***.*
2. *if the k-th column does not contain a leading nonzero entries, then* x_k *is a* **free variable***.*

The solution set of a linear system associated with a row echelon form can be written down by setting the free variables to be arbitrary constants and then solving for the leading variables using **backward induction**.

Example 2.1.5. *Write down the solution set of the linear systems associated with the given augmented matrices.*

1. $\left(\begin{array}{cccc|c} 1 & 3 & 0 & -2 & -3 \\ 0 & 1 & -5 & 1 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$
2. $\left(\begin{array}{cccc|c} 1 & -4 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$

Solution:

1. The only free variable is x_3. Let $x_3 = \alpha$. Then

$$\begin{aligned} x_4 &= 1 \\ x_2 &= 3 + 5x_3 - x_4 = 5\alpha + 2 \\ x_1 &= -3 - 3x_2 + 2x_4 = -15\alpha - 7 \end{aligned}$$

Thus the solution to the system is

$$(x_1, x_2, x_3, x_4) = (-15\alpha - 7, 5\alpha + 2, \alpha, 1)$$

where $\alpha \in \mathbb{R}$ is an arbitrary real number.

2. There are two free variables x_2 and x_3. Let $x_2 = \alpha$ and $x_3 = \beta$. Then

$$\begin{aligned} x_4 &= -3 \\ x_1 &= 2 + 4x_2 - 3x_3 = 4\alpha - 3\beta + 2 \end{aligned}$$

Thus the solution to the system is

$$(x_1, x_2, x_3, x_4) = (4\alpha - 3\beta + 2, \alpha, \beta, -3)$$

where $\alpha, \beta \in \mathbb{R}$ are arbitrary real numbers.

□

The solution set of the second example above is easier to be written down because it is in reduced row echelon form.

Definition 2.1.6 (Reduced row echelon form)**.** *A matrix* **R** *is said to be in* **reduced row echelon form** *(or* **R** *is a* **reduced row echelon matrix***) if it satisfies all the following properties:*

1. **R** *is in row echelon form.*
2. *Each leading nonzero entry of* **R** *is the only nonzero entry in its column.*

To solve a general linear system, we may transform the system to a row echelon form using elementary row operations.

Definition 2.1.7 (Elementary row operation)**.** *An* **elementary row operation** *is an operation on a matrix of one of the following forms.*

1. *Multiplying one row by a nonzero constant.*
2. *Interchanging two rows.*
3. *Replacing one row by adding a multiple of another row to it.*

Definition 2.1.8 (Row equivalent)**.** *We say that two matrices A and B are* **row equivalent** *if one can obtain B by applying successive elementary row operations to A.*

The following two theorems hold the keys of using Gaussian elimination to solve linear systems. The first theorem follows from the fact that elementary row operation does not alter the solution set.

Theorem 2.1.9. *Two linear systems have the same solution set if and only if the augmented matrices associated with the two systems are row equivalent.*

The second theorem says that any matrix is row equivalent to a row echelon form. It can be proved by induction on the number of columns.

Theorem 2.1.10. *Any matrix can be transformed to a row echelon form by applying successive elementary row operations.*

Let's see how Gaussian elimination works with the following examples.

Example 2.1.11. *Solve the linear system*

$$\left\{\begin{array}{rcrcrcr} x_1 & + & x_2 & - & 2x_3 & = & 5 \\ 2x_1 & - & x_2 & + & 2x_3 & = & -2 \\ x_1 & - & 2x_2 & + & 4x_3 & = & -4 \end{array}\right. .$$

Solution: The augmented matrix associated with the system is

$$\left(\begin{array}{ccc|c} 1 & 1 & -2 & 5 \\ 2 & -1 & 2 & -2 \\ 1 & -2 & 4 & -4 \end{array}\right)$$

Using Gaussian elimination, we have

$$\begin{aligned} & \left(\begin{array}{ccc|c} 1 & 1 & -2 & 5 \\ 2 & -1 & 2 & -2 \\ 1 & -2 & 4 & -4 \end{array}\right) \xrightarrow{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1}} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 5 \\ 0 & -3 & 6 & -12 \\ 0 & -3 & 6 & -9 \end{array}\right) \\ \xrightarrow{R_2 \to -\frac{1}{3}R_2} & \left(\begin{array}{ccc|c} 1 & 1 & -2 & 5 \\ 0 & 1 & -2 & 4 \\ 0 & -3 & 6 & -9 \end{array}\right) \xrightarrow{R_3 \to R_3 + 3R_2} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 5 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 0 & 3 \end{array}\right) \end{aligned}$$

The third row of the last matrix corresponds to the equation

$$0 = 3$$

which is absurd and thus has no solution. Therefore the solution set of the original linear system is empty. In other words, the linear system is inconsistent. □

Example 2.1.12. *Solve the linear system*

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 2 \\ x_1 & + & x_2 & + & x_3 & + & 2x_4 & + & 2x_5 & = & 3 \\ x_1 & + & x_2 & + & x_3 & + & 2x_4 & + & 3x_5 & = & 2 \end{array}\right. .$$

Solution: Using Gaussian elimination, we have

$$\begin{aligned} & \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right) \xrightarrow{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 \end{array}\right) \\ & \xrightarrow{R_3 \to R_3 - R_2} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right) \end{aligned}$$

Thus the system is equivalent to the following system

$$\left\{\begin{array}{rcrcrcrcrcr} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 2 \\ & & & & & & x_4 & + & x_5 & = & 1 \\ & & & & & & & & x_5 & = & -1 \end{array}\right. .$$

The solution of the system is

$$\begin{cases} x_5 = -1 \\ x_4 = 1 - x_5 = 2 \\ x_1 = 2 - x_2 - x_3 - x_4 - x_5 = 1 - x_2 - x_3 \end{cases}$$

Here x_1, x_4, x_5 are leading variables while x_2, x_3 are free variables. Another way of expressing the solution is

$$(x_1, x_2, x_3, x_4, x_5) = (1 - \alpha - \beta, \alpha, \beta, 2, -1),\ \ \alpha, \beta \in \mathbb{R}.$$

☐

A matrix may have many different row echelon forms. However, any matrix is row equivalent to a unique reduced row echelon form.

Theorem 2.1.13. *Every matrix is row equivalent to one and only one matrix in reduced row echelon form.*

Example 2.1.14. *Find the reduced row echelon form of the matrix*

$$\begin{pmatrix} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{pmatrix}.$$

Solution:

$$\begin{aligned}
&\begin{pmatrix} 1 & 2 & 1 & 4 \\ 3 & 8 & 7 & 20 \\ 2 & 7 & 9 & 23 \end{pmatrix} \xrightarrow{\substack{R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 - 2R_1}} \begin{pmatrix} 1 & 2 & 1 & 4 \\ 0 & 2 & 4 & 8 \\ 0 & 3 & 7 & 15 \end{pmatrix} \\
\xrightarrow{R_2 \to \frac{1}{2} R_2} &\begin{pmatrix} 1 & 2 & 1 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 7 & 15 \end{pmatrix} \xrightarrow{R_3 \to R_3 - 3R_2} \begin{pmatrix} 1 & 2 & 1 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix} \\
\xrightarrow{R_1 \to R_1 - 2R_2} &\begin{pmatrix} 1 & 0 & -3 & -4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix} \xrightarrow{\substack{R_1 \to R_1 + 3R_3 \\ R_2 \to R_2 - 2R_3}} \begin{pmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}
\end{aligned}$$

☐

Example 2.1.15. *Solve the linear system*

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 5 \\ x_1 + 2x_2 + 2x_3 + 3x_4 = 4 \\ x_1 + 2x_2 + x_3 + 2x_4 = 3 \end{cases}.$$

Solution:

$$\begin{aligned}
&\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 & 3 \end{array}\right) \xrightarrow{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -2 & -2 & -2 \end{array}\right) \\
\xrightarrow{R_2 \to -R_2} &\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 & -2 \end{array}\right) \xrightarrow{R_3 \to R_3 + 2R_2} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \\
\xrightarrow{R_1 \to R_1 - 3R_2} &\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)
\end{aligned}$$

Now x_1, x_3 are leading variables while x_2, x_4 are free variables. The solution of the system is

$$(x_1, x_2, x_3, x_4) = (2 - 2\alpha - \beta, \alpha, 1 - \beta, \beta), \ \alpha, \beta \in \mathbb{R}.$$

$\square$

Theorem 2.1.16. *Let* $\mathbf{R}$ *be a row echelon matrix which is row equivalent to the augmented matrix* $(\mathbf{A}|\mathbf{b})$ *of a linear system* $\mathbf{A}\mathbf{x}x = \mathbf{b}$*. Then the system has*

1. *no solution if the last column of* $\mathbf{R}$ *contains a leading nonzero entry.*
2. *unique solution if (1) does not holds and all variables are leading variables.*
3. *infinitely many solutions if (1) does not hold and there exists at least one free variable.*

Proof. Suppose $\mathbf{A}$ is an $m \times n$ matrix. Then $\mathbf{R}$ is an $m \times (n + 1)$ matrix. Since $\mathbf{R}$ and $(\mathbf{A}|\mathbf{b})$ are row equivalent, $\mathbf{R}$ has the same solution set with the system $\mathbf{A}\mathbf{x}x = \mathbf{b}$.

1. If the last column of $\mathbf{R}$ contains a leading nonzero entry, then the last nonzero row corresponds to the equation $0 = 1$ which has no solution.
2. If (1) does not hold and all variables are leading variables, then the value of x_k is uniquely determined by the equation associated with the k-th row.
3. If (1) does not hold and there exists at least one free variable, then there are infinitely many possible values for the free variable.

$\square$

In particular, the homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{0}$ always has the solution $\mathbf{x} = \mathbf{0}$. This solution is called the **trivial solution**. The trivial solution is the only solution to the homogeneous system if and only if $\mathbf{A}$ is row equivalent to I.

Theorem 2.1.17. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. Then homogeneous linear system* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *has only the trivial solution if and only if* $\mathbf{A}$ *is row equivalent to the identity matrix* $\mathbf{I}$*.*

Proof. The system $\mathbf{A}\mathbf{x} = \mathbf{0}$ has at least one solution namely the trivial solution $\mathbf{x} = \mathbf{0}$. Let $\mathbf{R}$ be the reduced row echelon form of $(\mathbf{A}|\mathbf{0})$. By Theorem 2.1.16, the trivial solution is the only solution if and only if there is no free variable. This is the case if and only if $\mathbf{R} = (\mathbf{I}|\mathbf{0})$ or equivalently $\mathbf{A}$ is row equivalent to $\mathbf{I}$. $\square$

Exercise 2.1

1. Find the reduced row echelon form of the following matrices.

   (a) $\begin{pmatrix} 3 & 7 & 15 \\ 2 & 5 & 11 \end{pmatrix}$

   (b) $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 2 & 1 & 9 \end{pmatrix}$

   (c) $\begin{pmatrix} 5 & 2 & -5 \\ 9 & 4 & -7 \\ 4 & 1 & -7 \end{pmatrix}$

   (d) $\begin{pmatrix} 1 & -4 & -2 \\ 3 & -12 & 1 \\ 2 & -8 & 5 \end{pmatrix}$

(e) $\begin{pmatrix} 2 & 2 & 4 & 2 \\ 1 & -1 & -4 & 3 \\ 2 & 7 & 19 & -3 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & -2 & -4 & 5 \\ -2 & 4 & -3 & 1 \\ 3 & -6 & -1 & 4 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 2 & 3 & 4 \\ -1 & -2 & -1 & -2 & -3 \end{pmatrix}$

(h) $\begin{pmatrix} 3 & 6 & 1 & 7 & 13 \\ 5 & 10 & 8 & 18 & 47 \\ 2 & 4 & 5 & 9 & 26 \end{pmatrix}$

(i) $\begin{pmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{pmatrix}$

2. Solve the following systems of linear equations.

(a) $\left\{\begin{array}{rcrcrcr} x_1 & - & 3x_2 & + & 4x_3 & = & 7 \\ & & x_2 & - & 5x_3 & = & 2 \end{array}\right.$

(b) $\left\{\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & 3x_3 & = & -4 \\ x_1 & + & x_2 & + & x_3 & = & 1 \\ 5x_1 & + & 6x_2 & + & 8x_3 & = & 8 \end{array}\right.$

(c) $\left\{\begin{array}{rcrcrcr} 2x_1 & - & x_2 & + & 5x_3 & = & 15 \\ x_1 & + & 3x_2 & - & x_3 & = & 4 \\ x_1 & - & 4x_2 & + & 6x_3 & = & 11 \\ 3x_1 & + & 9x_2 & - & 3x_3 & = & 12 \end{array}\right.$

(d) $\left\{\begin{array}{rcrcrcrcr} x_1 & + & x_2 & - & 2x_3 & + & x_4 & = & 9 \\ & & x_2 & - & x_3 & + & 2x_4 & = & 1 \\ & & & & x_3 & - & 3x_4 & = & 5 \end{array}\right.$

(e) $\left\{\begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & - & 2x_2 & + & x_3 & - & x_4 & = & -1 \\ x_1 & - & 2x_2 & + & x_3 & + & 5x_4 & = & 5 \end{array}\right.$

(f) $\left\{\begin{array}{rcrcrcrcr} 3x_1 & - & 6x_2 & + & x_3 & + & 13x_4 & = & 15 \\ 3x_1 & - & 6x_2 & + & 3x_3 & + & 21x_4 & = & 21 \\ 2x_1 & - & 4x_2 & + & 5x_3 & + & 26x_4 & = & 23 \end{array}\right.$

2.2 Matrix arithmetic

A **matrix** is a rectangular array of real (or complex) numbers of the form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

The horizontal arrays are called **rows** and the vertical arrays are called **columns**. There are m rows and n columns in the above matrix and it is called an $m \times n$ matrix. We called the number in the i-th row and j-th column, where is a_{ij} in the above matrix, the ij-th **entry** of the matrix. If the number of rows of a matrix is equal to the number of its columns, then it is called a **square matrix**.

Definition 2.2.1. *The arithmetic of matrices are defined as follows.*

1. **Addition**: *Let* $\mathbf{A} = [a_{ij}]$ *and* $\mathbf{B} = [b_{ij}]$ *be two* $m \times n$ *matrices. Then*

$$[\mathbf{A} + \mathbf{B}]_{ij} = a_{ij} + b_{ij}.$$

That is

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

2. **Scalar multiplication**: *Let* $\mathbf{A} = [a_{ij}]$ *be an* $m \times n$ *matrix and* c *be a scalar. Then*

$$[c\mathbf{A}]_{ij} = ca_{ij}.$$

That is

$$c\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}.$$

3. **Matrix multiplication**: *Let* $\mathbf{A} = [a_{ij}]$ *be an* $m \times n$ *matrix and* $\mathbf{B} = [b_{jk}]$ *be an* $n \times r$*. Then their matrix product* $\mathbf{AB}$ *is an* $m \times r$ *matrix where its* ik*-th entry is*

$$[\mathbf{AB}]_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}.$$

For example: If $\mathbf{A}$ *is a* 3×2 *matrix and* $\mathbf{B}$ *is a* 2×2 *matrix, then*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} \end{pmatrix}$$

is a 3×2 *matrix.*

A **zero matrix**, denoted by $\mathbf{0}$, is a matrix whose entries are all zeros. An **identity matrix**, denoted by $\mathbf{I}$, is a square matrix that has ones on its principal diagonal and zero elsewhere.

Theorem 2.2.2 (Properties of matrix algebra)**.** *Let* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *be matrices of appropriate sizes to make the indicated operations possible and* a, b *be real numbers, then following identities hold.*

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
3. $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$
4. $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$

5. $(a+b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$
6. $a(b\mathbf{A}) = (ab)\mathbf{A}$
7. $a(\mathbf{AB}) = (a\mathbf{A})\mathbf{B} = \mathbf{A}(a\mathbf{B})$
8. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
9. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
10. $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
11. $\mathbf{A0} = \mathbf{0A} = \mathbf{0}$
12. $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$

Proof. All properties are obvious except (8) and we prove it here. Let $\mathbf{A} = [a_{ij}]$ be $m \times n$ matrix, $\mathbf{B} = [b_{jk}]$ be $n \times r$ matrix and $\mathbf{C} = [c_{kl}]$ be $r \times s$ matrix. Then

$$\begin{aligned}
[(\mathbf{AB})\mathbf{C}]_{il} &= \sum_{k=1}^{r} [\mathbf{AB}]_{ik} c_{kl} \\
&= \sum_{k=1}^{r} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) c_{kl} \\
&= \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{r} b_{jk} c_{kl} \right) \\
&= \sum_{j=1}^{n} a_{ij} [\mathbf{BC}]_{jl} \\
&= [\mathbf{A}(\mathbf{BC})]_{il}
\end{aligned}$$

□

Remarks:

1. $\mathbf{AB}$ is defined only when the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$.

2. In general, $\mathbf{AB} \neq \mathbf{BA}$ even when they are both well-defined and of the same type. For example:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

   Then

$$\mathbf{AB} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$

$$\mathbf{BA} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

3. $\mathbf{AB} = \mathbf{0}$ does not implies that $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$. For example:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \mathbf{0} \text{ and } \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \neq \mathbf{0}$$

   But

$$\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Definition 2.2.3. *Let* $\mathbf{A} = [a_{ij}]$ *be an* $m \times n$ *matrix. Then the* **transpose** *of* $\mathbf{A}$ *is the* $n \times m$ *matrix defined by interchanging rows and columns and is denoted by* $\mathbf{A}^T$*, i.e.,*

$$[\mathbf{A}^T]_{ji} = a_{ij} \textit{ for } 1 \leq j \leq n, 1 \leq i \leq m.$$

Example 2.2.4.

1. $\begin{pmatrix} 2 & 0 & 5 \\ 4 & -1 & 7 \end{pmatrix}^T = \begin{pmatrix} 2 & 4 \\ 0 & -1 \\ 5 & 7 \end{pmatrix}$

2. $\begin{pmatrix} 7 & -2 & 6 \\ 1 & 2 & 3 \\ 5 & 0 & 4 \end{pmatrix}^T = \begin{pmatrix} 7 & 1 & 5 \\ -2 & 2 & 0 \\ 6 & 3 & 4 \end{pmatrix}$

□

Theorem 2.2.5 (Properties of transpose)**.** *For any* $m \times n$ *matrices* $\mathbf{A}$ *and* $\mathbf{B}$*,*

1. $(\mathbf{A}^T)^T = \mathbf{A}$*;*

2. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$*;*

3. $(c\mathbf{A})^T = c\mathbf{A}^T$*;*

4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

Definition 2.2.6 (Symmetric and skew-symmetric matrices)**.** *Let* $\mathbf{A}$ *be a square matrix.*

1. We say that $\mathbf{A}$ *is* **symmetric** *if* $\mathbf{A}^T = \mathbf{A}$.

2. We say that $\mathbf{A}$ *is* **anti-symmetric** *(or skew-symmetric) if* $\mathbf{A}^T = -\mathbf{A}$.

Exercise 2.2

1. Find a 2×2 matrix $\mathbf{A}$ such that $\mathbf{A}^2 = \mathbf{0}$ but $\mathbf{A} \neq \mathbf{0}$.

2. Find a 2×2 matrix $\mathbf{A}$ such that $\mathbf{A}^2 = \mathbf{I}$ but $\mathbf{A} \neq \pm\mathbf{I}$.

3. Let $\mathbf{A}$ be a square matrix. Prove that $\mathbf{A}$ can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

4. Suppose $\mathbf{A}$, $\mathbf{B}$ are symmetric matrices and $\mathbf{C}$, $\mathbf{D}$ are skew-symmetric matrices such that $\mathbf{A} + \mathbf{C} = \mathbf{B} + \mathbf{D}$. Prove that $\mathbf{A} = \mathbf{B}$ and $\mathbf{C} = \mathbf{D}$

5. Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

   Prove that

$$\mathbf{A}^2 - (a + d)\mathbf{A} + (ad - bc)\mathbf{I} = \mathbf{0}$$

6. Let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ matrices. Prove that $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$ if and only if $\mathbf{AB} = \mathbf{BA}$.

2.3 Inverse

For square matrices, we have an important notion of inverse matrix.

Definition 2.3.1 (Inverse)**.** *A square matrix* $\mathbf{A}$ *is said to be* **invertible**, *if there exists a matrix* $\mathbf{B}$ *such that*

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}.$$

We say that $\mathbf{B}$ *is a (multiplicative)* **inverse** *of* A.

Inverse of a matrix is unique if it exists. Thus it makes sense to say the inverse of a square matrix.

Theorem 2.3.2. *If* $\mathbf{A}$ *is invertible, then the inverse of* A *is unique.*

Proof. Suppose $\mathbf{B}_1$ and $\mathbf{B}_2$ are multiplicative inverses of $\mathbf{A}$. Then

$$\mathbf{B}_2 = \mathbf{B}_2\mathbf{I} = \mathbf{B}_2(\mathbf{AB}_1) = (\mathbf{B}_2\mathbf{A})\mathbf{B}_1 = \mathbf{IB}_1 = \mathbf{B}_1.$$

□

The unique inverse of an invertible matrix $\mathbf{A}$ is denoted by $\mathbf{A}^{-1}$.

Example 2.3.3. *The* 2×2 *matrix*

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible if and only if $ad - bc \neq 0$, *in which case*

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Theorem 2.3.4. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two invertible* $n \times n$ *matrices.*

1. $\mathbf{A}^{-1}$ *is invertible and* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$;
2. *For any nonnegative integer* k, $\mathbf{A}^k$ *is invertible and* $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$;
3. *The product* $\mathbf{AB}$ *is invertible and*
$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1};$$
4. $\mathbf{A}^T$ *is invertible and*
$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T.$$

Proof. We prove (3) only.

$$\begin{aligned}(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) &= \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I} \\ (\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) &= \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{IB} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}\end{aligned}$$

Therefore $\mathbf{AB}$ is invertible and $\mathbf{B}^{-1}\mathbf{A}^{-1}$ is the inverse of $\mathbf{AB}$. □

Inverse matrix can be used to solve linear system.

Theorem 2.3.5. *If the* $n \times n$ *matrix* $\mathbf{A}$ *is invertible, then for any* n*-vector* $\mathbf{b}$ *the system* $\mathbf{Ax} = \mathbf{b}$ *has the unique solution* $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Example 2.3.6. *Solve the system*

$$\begin{cases} 4x_1 &+& 6x_2 &=& 6 \\ 5x_1 &+& 9x_2 &=& 18 \end{cases}.$$

Solution: Let $\mathbf{A} = \begin{pmatrix} 4 & 6 \\ 5 & 9 \end{pmatrix}$. Then

$$\mathbf{A}^{-1} = \frac{1}{6}\begin{pmatrix} 9 & -6 \\ -5 & 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & -1 \\ -\frac{5}{6} & \frac{2}{3} \end{pmatrix}$$

Thus the solution is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \begin{pmatrix} \frac{3}{2} & -1 \\ -\frac{5}{6} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} 6 \\ 18 \end{pmatrix} = \begin{pmatrix} -9 \\ 7 \end{pmatrix}$$

Therefore $(x_1, x_2) = (-9, 7)$. □

Next we discuss how to find the inverse of an invertible matrix.

Definition 2.3.7. *A square matrix* **E** *is called an* **elementary matrix** *if it can be obtained by performing a single elementary row operation on* **I**.

The relationship between elementary row operation and elementary matrix is given in the following theorem which can be proved easily case by case.

Theorem 2.3.8. *Let* **E** *be the elementary matrix obtained by performing a certain elementary row operation on* **I**. *Then the result of performing the same elementary row operation on a matrix* **A** *is* **EA**.

Theorem 2.3.9. *Every elementary matrix is invertible.*

The above theorem can also by proved case by case. In stead of giving a rigorous proof, let's look at some examples.

Example 2.3.10. *Examples of elementary matrices associated to elementary row operations and their inverses.*

Elementary row operation	*Interchanging two rows*	*Multiplying a row by a nonzero constant*	*Adding a multiple of a row to another row*
Elementary matrix	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Inverse	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Theorem 2.3.11. *Let* **A** *be a square matrix. Then the following statements are equivalent.*

1. **A** *is invertible*

2. $\mathbf{A}$ *is row equivalent to* $\mathbf{I}$
3. $\mathbf{A}$ *is a product of elementary matrices*

Proof. The theorem follows easily from the fact that an $n \times n$ reduced row echelon matrix is invertible if and only if it is the identity matrix $\mathbf{I}$. □

Let $\mathbf{A}$ be an invertible matrix. Then the above theorem tells us that there exists elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_k$ such that

$$\mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{I}.$$

Multiplying both sides by $(\mathbf{E}_1)^{-1}(\mathbf{E}_2)^{-1} \cdots (\mathbf{E}_{k-1})^{-1}(\mathbf{E}_k)^{-1}$ we have

$$\mathbf{A} = (\mathbf{E}_1)^{-1}(\mathbf{E}_2)^{-1} \cdots (\mathbf{E}_{k-1})^{-1}(\mathbf{E}_k)^{-1}.$$

Therefore

$$\mathbf{A}^{-1} = \mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_2 \mathbf{E}_1$$

by Proposition 2.3.4.

Theorem 2.3.12. *Let* $\mathbf{A}$ *be a square matrix. Suppose we can preform elementary row operation to the augmented matrix* $(\mathbf{A}|\mathbf{I})$ *to obtain a matrix of the form* $(\mathbf{I}|\mathbf{E})$*, then* $\mathbf{A}^{-1} = \mathbf{E}$.

Proof. Let $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_k$ be elementary matrices such that

$$\mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_2 \mathbf{E}_1 (\mathbf{A}|\mathbf{I}) = (\mathbf{I}|\mathbf{E}).$$

Then the multiplication on the left submatrix gives

$$\mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{I}$$

and the multiplication of the right submatrix gives

$$\mathbf{E} = \mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{I} = \mathbf{A}^{-1}.$$

□

Example 2.3.13. *Find the inverse of*

$$\begin{pmatrix} 4 & 3 & 2 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix}$$

Solution:

$$\left(\begin{array}{ccc|ccc} 4 & 3 & 2 & 1 & 0 & 0 \\ 5 & 6 & 3 & 0 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_1 \to R_1 - R_3} \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & -1 \\ 5 & 6 & 3 & 0 & 1 & 0 \\ 3 & 5 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{\substack{R_2 \to R_2 - 5R_1 \\ R_3 \to R_3 - 3R_1}} \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & -1 \\ 0 & 16 & 3 & -5 & 1 & 5 \\ 0 & 11 & 2 & -3 & 0 & 4 \end{array}\right) \xrightarrow{R_2 \to R_2 - R_3} \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & -1 \\ 0 & 5 & 1 & -2 & 1 & 1 \\ 0 & 11 & 2 & -3 & 0 & 4 \end{array}\right)$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & -1 \\ 0 & 5 & 1 & -2 & 1 & 1 \\ 0 & 1 & 0 & 1 & -2 & 2 \end{array}\right) \xrightarrow{R_2 \leftrightarrow R_3} \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 5 & 1 & -2 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{R_3 \to R_3 - 5R_2} \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -7 & 11 & -9 \end{array}\right) \xrightarrow{R_1 \to R_1 + 2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -4 & 3 \\ 0 & 1 & 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -7 & 11 & -9 \end{array}\right)$$

Therefore

$$\mathbf{A}^{-1} = \begin{pmatrix} 3 & -4 & 3 \\ 1 & -2 & 2 \\ -7 & 11 & -9 \end{pmatrix}.$$

□

Example 2.3.14. *Find a* 3×2 *matrix* $\mathbf{X}$ *such that*

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix} \mathbf{X} = \begin{pmatrix} 0 & -3 \\ -1 & 4 \\ 2 & 1 \end{pmatrix}.$$

Solution:

$$\begin{aligned}
&\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & -3 \\ 2 & 1 & 2 & -1 & 4 \\ 1 & 3 & 4 & 2 & 1 \end{array}\right)
&& \xrightarrow{\substack{R_2 \to R_2 - 5R_1 \\ R_3 \to R_3 - 3R_1}}
\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & -3 \\ 0 & -3 & -4 & -1 & 10 \\ 0 & 1 & 1 & 2 & 4 \end{array}\right) \\
\xrightarrow{R_2 \leftrightarrow R_3} &\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & -3 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -3 & -4 & -1 & 10 \end{array}\right)
&& \xrightarrow{R_3 \to R_3 + 3R_2}
\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & -3 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & 0 & -1 & 5 & 22 \end{array}\right) \\
\xrightarrow{R_3 \to -R_3} &\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & -3 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & 0 & 1 & -5 & -22 \end{array}\right)
&& \xrightarrow{\substack{R_1 \to R_1 - 3R_3 \\ R_2 \to R_2 - R_3}}
\left(\begin{array}{ccc|cc} 1 & 2 & 0 & 15 & 63 \\ 0 & 1 & 0 & 7 & 26 \\ 0 & 5 & 1 & -5 & -22 \end{array}\right) \\
\xrightarrow{R_1 \to R_1 - 2R_2} &\left(\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 11 \\ 0 & 1 & 0 & 7 & 26 \\ 0 & 0 & 1 & -5 & -22 \end{array}\right)
\end{aligned}$$

Therefore we may take

$$\mathbf{X} = \begin{pmatrix} 1 & 11 \\ 7 & 26 \\ -5 & -22 \end{pmatrix}.$$

□

Exercise 2.3

1. Find the inverse of the following matrices.

   (a) $\begin{pmatrix} 5 & 6 \\ 4 & 5 \end{pmatrix}$

   (b) $\begin{pmatrix} 5 & 7 \\ 4 & 6 \end{pmatrix}$

   (c) $\begin{pmatrix} 1 & 5 & 1 \\ 2 & 5 & 0 \\ 2 & 7 & 1 \end{pmatrix}$

   (d) $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 8 & 3 \\ 3 & 10 & 6 \end{pmatrix}$

   (e) $\begin{pmatrix} 1 & -3 & -3 \\ -1 & 1 & 2 \\ 2 & -3 & -3 \end{pmatrix}$

   (f) $\begin{pmatrix} 1 & -2 & 2 \\ 3 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$

   (g) $\begin{pmatrix} 4 & 0 & 1 & 1 \\ 3 & 1 & 3 & 1 \\ 0 & 1 & 2 & 0 \\ 3 & 2 & 4 & 1 \end{pmatrix}$

2. Solve the following systems of equations by finding the inverse of the coefficient matrices.

(a) $\left\{\begin{array}{rcrcr} x_1 & + & x_2 & = & 2 \\ 5x_1 & + & 6x_2 & = & 9 \end{array}\right.$. (b) $\left\{\begin{array}{rcrcrcr} 5x_1 & + & 3x_2 & + & 2x_3 & = & 4 \\ 3x_1 & + & 3x_2 & + & 2x_3 & = & 2 \\ & & x_2 & + & x_3 & = & 5 \end{array}\right.$.

3. Solve the following matrix equations for $\mathbf{X}$.

(a) $\begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & 1 \\ 2 & -1 & 4 \end{pmatrix} \mathbf{X} = \begin{pmatrix} 2 & 1 \\ 2 & 0 \\ 3 & 5 \end{pmatrix}$. (b) $\begin{pmatrix} 1 & -1 & 1 \\ 2 & 3 & 0 \\ 0 & 2 & -1 \end{pmatrix} \mathbf{X} = \begin{pmatrix} -1 & 5 \\ 0 & -3 \\ 5 & -7 \end{pmatrix}$.

4. Suppose $\mathbf{A}$ is an invertible matrix and $\mathbf{B}$ is a matrix such that $\mathbf{A}+\mathbf{B}$ is invertible. Show that $\mathbf{A}(\mathbf{A}+\mathbf{B})^{-1}$ is the inverse of $\mathbf{I}+\mathbf{B}\mathbf{A}^{-1}$.

5. Let $\mathbf{A}$ be a square matrix such that $\mathbf{A}^k = \mathbf{0}$ for some positive integer k. Show that $\mathbf{I}-\mathbf{A}$ is invertible.

6. Show that if $\mathbf{A}$ and $\mathbf{B}$ are invertible matrices such that $\mathbf{A}+\mathbf{B}$ is invertible, then $\mathbf{A}^{-1}+\mathbf{B}^{-1}$ is invertible.

7. Let $\mathbf{A}(t)$ be a matrix valued function such that all entries are differentiable functions of t and $\mathbf{A}(t)$ is invertible for any t. Prove that

$$\frac{d}{dt}\left(\mathbf{A}^{-1}\right) = -\mathbf{A}^{-1}\left(\frac{d}{dt}\mathbf{A}\right)\mathbf{A}^{-1}$$

8. Suppose $\mathbf{A}$ is a square matrix such that there exists non-singular symmetric matrix[1] with $\mathbf{A}+\mathbf{A}^T = \mathbf{S}^2$. Prove that $\mathbf{A}$ is non-singular.

2.4 Determinant

We can associate an important quantity called determinant to every square matrix. The determinant of a square matrix has enormous meanings. For example if the its value is non-zero, then the matrix is invertible, the homogeneous system associated with the matrix does not have non-trivial solution and the row vectors, or column vectors of the matrix are linearly independent.

The determinant of a square matrix can be defined inductively. The determinant of a 1×1 matrix is the value of the its only entry. Suppose we have defined the determinant of an $(n-1)\times(n-1)$ matrix. Then the determinant of an $n\times n$ matrix is defined in terms of its cofactors.

Definition 2.4.1 (Minor and cofactor)**.** *Let $\mathbf{A} = [a_{ij}]$ be an $n \times n$ matrix.*

1. *The ij-th* **minor** *of $\mathbf{A}$ is the determinant M_{ij} of the $(n-1)\times(n-1)$ submatrix that remains after deleting the i-th row and the j-th column of $\mathbf{A}$.*
2. *The ij-th* **cofactor** *of $\mathbf{A}$ is defined by*

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

As we can see, when defining cofactors of an $n\times n$ matrix, the determinants of $(n-1)\times(n-1)$ matrices are involved. Now we can use cofactor to define the determinant of an $n\times n$ matrix inductively.

---

[1] A square matrix $\mathbf{S}$ is symmetric if $\mathbf{S}^T = \mathbf{S}$.

Definition 2.4.2 (Determinant)**.** *Let* $\mathbf{A} = [a_{ij}]$ *be an* $n \times n$ *matrix. The* **determinant** $\det(\mathbf{A})$ *of* $\mathbf{A}$ *is defined inductively as follows.*

1. *If* $n = 1$, *then* $\det(\mathbf{A}) = a_{11}$.
2. *If* $n > 1$, *then*

$$\det(\mathbf{A}) = \sum_{k=1}^{n} a_{1k}A_{1k} = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n},$$

where A_{ij} *is the* ij*-th cofactor of* $\mathbf{A}$.

Example 2.4.3. *When* $n = 1, 2$ *or* 3, *we have the following.*

1. *The determinant of a* 1×1 *matrix is*

$$|a_{11}| = a_{11}$$

2. *The determinant of a* 2×2 *matrix is*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

3. *The determinant of a* 3×3 *matrix is*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Example 2.4.4.

$$\begin{aligned}
& \begin{vmatrix} 4 & 3 & 0 & 1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 1 \end{vmatrix} \\
= \; & 4\begin{vmatrix} 2 & 0 & 1 \\ 0 & 0 & 3 \\ 1 & 2 & 1 \end{vmatrix} - 3\begin{vmatrix} 3 & 0 & 1 \\ 1 & 0 & 3 \\ 0 & 2 & 1 \end{vmatrix} + 0\begin{vmatrix} 3 & 2 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{vmatrix} - 1\begin{vmatrix} 3 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{vmatrix} \\
= \; & 4\left(2\begin{vmatrix} 0 & 3 \\ 2 & 1 \end{vmatrix} - 0\begin{vmatrix} 0 & 3 \\ 1 & 1 \end{vmatrix} + 1\begin{vmatrix} 0 & 0 \\ 1 & 2 \end{vmatrix}\right) \\
& -3\left(3\begin{vmatrix} 0 & 3 \\ 2 & 1 \end{vmatrix} - 0\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} + 1\begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix}\right) \\
& -\left(3\begin{vmatrix} 0 & 0 \\ 1 & 2 \end{vmatrix} - 2\begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} + 0\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\right) \\
= \; & 4\,(2(-6)) - 3\,(3(-6) + 1(2)) - (-2(2)) \\
= \; & 4
\end{aligned}$$

□

The following theorem can be proved by induction on n.

Theorem 2.4.5. *Let $\mathbf{A} = [a_{ij}]$ be an $n \times n$ matrix. Then*

$$\det(\mathbf{A}) = \sum_{\sigma \in S_n} sign(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

where S_n is the set of all permutations[2] *of $\{1, 2, \cdots, n\}$ and*

$$sign(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is an even permutation,} \\ -1 & \text{if } \sigma \text{ is an odd permutation.} \end{cases}$$

Note that there are $n!$ number of terms for an $n \times n$ determinant in the above formula. Here we write down the $4! = 24$ terms of a 4×4 determinant.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = \begin{array}{l} a_{11}a_{22}a_{33}a_{44} - a_{11}a_{22}a_{34}a_{43} - a_{11}a_{23}a_{32}a_{44} + a_{11}a_{23}a_{34}a_{42} \\ +a_{11}a_{24}a_{32}a_{43} - a_{11}a_{24}a_{33}a_{42} - a_{12}a_{21}a_{33}a_{44} + a_{12}a_{21}a_{34}a_{43} \\ +a_{12}a_{23}a_{31}a_{44} - a_{12}a_{23}a_{34}a_{41} - a_{12}a_{24}a_{31}a_{43} + a_{12}a_{24}a_{33}a_{41} \\ +a_{13}a_{21}a_{32}a_{44} - a_{13}a_{21}a_{34}a_{42} - a_{13}a_{22}a_{31}a_{44} + a_{13}a_{22}a_{34}a_{41} \\ +a_{13}a_{24}a_{31}a_{42} - a_{13}a_{24}a_{32}a_{41} - a_{14}a_{21}a_{32}a_{43} + a_{14}a_{21}a_{33}a_{42} \\ +a_{14}a_{22}a_{31}a_{43} - a_{14}a_{22}a_{33}a_{41} - a_{14}a_{23}a_{31}a_{42} + a_{14}a_{23}a_{32}a_{41} \end{array}$$

By Theorem 2.4.5, it is easy to see the following.

Theorem 2.4.6. *The determinant of an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ can be obtained by expansion along any row or column, i.e., for any $1 \le i \le n$, we have*

$$\det(\mathbf{A}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$$

and for any $1 \le j \le n$, we have

$$\det(\mathbf{A}) = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}.$$

Example 2.4.7. *We can expand the determinant along the 3rd column in Example 2.4.4.*

$$\begin{array}{rl} & \begin{vmatrix} 4 & 3 & 0 & 1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 1 \end{vmatrix} \\ = & -2 \begin{vmatrix} 4 & 3 & 1 \\ 3 & 2 & 1 \\ 1 & 0 & 3 \end{vmatrix} \\ = & -2 \left(-3 \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} + 2 \begin{vmatrix} 4 & 1 \\ 1 & 3 \end{vmatrix} \right) \\ = & -2\,(-3(8) + 2(11)) \\ = & 4 \end{array}$$

□

Theorem 2.4.8. *Properties of determinant.*

1. $\det(\mathbf{I}) = 1$;

---

[2] A transposition is a permutation which swaps two numbers and keep all other fixed. A permutation is even, odd if it is a composition of an even, odd number of transpositions respectively.

2. *Suppose that the matrices* $\mathbf{A}_1$, $\mathbf{A}_2$ *and* $\mathbf{B}$ *are identical except for their i-th row (or column) and that the i-th row (or column) of* $\mathbf{B}$ *is the sum of the i-th row (or column) of* $\mathbf{A}_1$ *and* $\mathbf{A}_2$, *then* $\det(\mathbf{B}) = \det(\mathbf{A}_1) + \det(\mathbf{A}_2)$;
3. *If* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by multiplying a single row (or column) of* $\mathbf{A}$ *by the constant* k, *then* $\det(\mathbf{B}) = k\det(\mathbf{A})$;
4. *If* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by interchanging two rows (or columns), then* $\det(\mathbf{B}) = -\det(\mathbf{A})$;
5. *If* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by adding a constant multiple of one row (or column) of* $\mathbf{A}$ *to another row (or column) of* $\mathbf{A}$, *then* $\det(\mathbf{B}) = \det(\mathbf{A})$;
6. *If two rows (or columns) of* $\mathbf{A}$ *are identical, then* $\det(\mathbf{A}) = 0$;
7. *If* $\mathbf{A}$ *has a row (or column) consisting entirely of zeros, then* $\det(\mathbf{A}) = 0$;
8. $\det(\mathbf{A}^T) = \det(\mathbf{A})$;
9. *If* $\mathbf{A}$ *is a triangular matrix, then* $\det(\mathbf{A})$ *is the product of the diagonal elements of* $\mathbf{A}$;
10. $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.

All the statements in the above theorem are simple consequences of Theorem 2.4.5 or Theorem 2.4.6 except (10) which will be proved later in this section (Theorem 2.4.16). Statements (3), (4), (5) allow us to evaluate a determinant using row or column operations.

Example 2.4.9.

$$\begin{aligned}
&\begin{vmatrix} 2 & 2 & 5 & 5 \\ 1 & -2 & 4 & 1 \\ -1 & 2 & -2 & -2 \\ -2 & 7 & -3 & 2 \end{vmatrix} \\
= &\begin{vmatrix} 0 & 6 & -3 & 3 \\ 1 & -2 & 4 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 3 & 5 & 4 \end{vmatrix} \quad \begin{pmatrix} R_1 \to R_1 - 2R_2 \\ R_3 \to R_3 + R_2 \\ R_4 \to R_4 + 2R_2 \end{pmatrix} \\
= & -\begin{vmatrix} 6 & -3 & 3 \\ 0 & 2 & -1 \\ 3 & 5 & 4 \end{vmatrix} \\
= & -3\begin{vmatrix} 2 & -1 & 1 \\ 0 & 2 & -1 \\ 3 & 5 & 4 \end{vmatrix} \\
= & -3\left(2\begin{vmatrix} -1 & 1 \\ 5 & 4 \end{vmatrix} + 3\begin{vmatrix} -1 & 1 \\ 2 & -1 \end{vmatrix}\right) \\
= & -69
\end{aligned}$$

□

We can also use column operations.

Example 2.4.10.

$$\begin{aligned}
&\begin{vmatrix} 2 & 2 & 5 & 5 \\ 1 & -2 & 4 & 1 \\ -1 & 2 & -2 & -2 \\ -2 & 7 & -3 & 2 \end{vmatrix} \\
= &\begin{vmatrix} 2 & 6 & -3 & 3 \\ 1 & 0 & 0 & 0 \\ -1 & 0 & 2 & -1 \\ -2 & 3 & 5 & 4 \end{vmatrix} \begin{pmatrix} C_2 \to C_2 + 2C_1 \\ C_3 \to C_3 - 4C_1 \\ C_4 \to C_4 - C_1 \end{pmatrix} \\
= & -\begin{vmatrix} 6 & -3 & 3 \\ 0 & 2 & -1 \\ 3 & 5 & 4 \end{vmatrix} \\
= & -\begin{vmatrix} 0 & 0 & 3 \\ 2 & 1 & -1 \\ -5 & 9 & 4 \end{vmatrix} \begin{pmatrix} C_1 \to C_1 - 2C_3 \\ C_2 \to C_2 + C_3 \end{pmatrix} \\
= & -3\begin{vmatrix} 2 & 1 \\ -5 & 9 \end{vmatrix} \\
= & -69
\end{aligned}$$

□

Some determinants can be evaluated using the properties of determinants.

Example 2.4.11. *Let* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *be real numbers and*

$$\mathbf{A} = \begin{pmatrix} 1 & \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ 1 & x & \alpha_2 & \cdots & \alpha_n \\ 1 & \alpha_1 & x & \cdots & \alpha_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_1 & \alpha_2 & \cdots & x \end{pmatrix}.$$

Show that

$$\det(\mathbf{A}) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).$$

Solution: Note that A is an $(n+1) \times (n+1)$ matrix. For simplicity we assume that $\alpha_1, \alpha_2, \cdots, \alpha_n$ are distinct. Observe that we have the following 3 facts.

1. $\det(\mathbf{A})$ is a polynomial of degree n in x;
2. $\det(\mathbf{A}) = 0$ when $x = \alpha_i$ for some i;
3. The coefficient of x^n of $\det(\mathbf{A})$ is 1.

Then the equality follows by the factor theorem. □

Example 2.4.12. *The Vandermonde determinant is defined as*

$$V(x_1, x_2, \cdots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}.$$

Show that

$$V(x_1, x_2, \cdots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Solution: Using factor theorem, the equality is a consequence of the following 3 facts.

1. $V(x_1, x_2, \cdots, x_n)$ is a polynomial of degree $n(n-1)/2$ in $x_1, x_2, \cdots, x_n$;
2. For any $i \neq j$, $V(x_1, x_2, \cdots, x_n) = 0$ when $x_i = x_j$;
3. The coefficient of $x_2 x_3^2 \cdots x_n^{n-1}$ of $V(x_1, x_2, \cdots, x_n)$ is 1.

□

Now we are going to prove (10) of Theorem 2.4.8. The following lemma says that the statement is true when one of the matrix is an elementary matrix.

Lemma 2.4.13. *Let* $\mathbf{A} = [a_{ij}]$ *be an* $n \times n$ *matrix and* $\mathbf{E}$ *be an* $n \times n$ *elementary matrix. Then*

$$\det(\mathbf{EA}) = \det(\mathbf{E}) \det(\mathbf{A}).$$

Proof. The statement can be checked for each of the 3 types of elementary matrix E. □

Definition 2.4.14. *Let* $\mathbf{A}$ *be a square matrix. We say that* $\mathbf{A}$ *is* **singular** *if the system* $\mathbf{Ax} = \mathbf{0}$ *has non-trivial solution. A square matrix is* **non-singular** *if it is not singular.*

Theorem 2.4.15. *The following conditions for an* $n \times n$ *matrix* $\mathbf{A}$ *are equivalent.*

1. $\mathbf{A}$ *is non-singular, i.e., the system* $\mathbf{Ax} = \mathbf{0}$ *has only trivial solution* $\mathbf{x} = \mathbf{0}$.
2. $\mathbf{A}$ *is invertible, i.e.,* $\mathbf{A}^{-1}$ *exists.*
3. $\det(\mathbf{A}) \neq 0$.
4. $\mathbf{A}$ *is row equivalent to* $\mathbf{I}$.
5. *For any n-column vector* $\mathbf{b}$*, the system* $\mathbf{Ax} = \mathbf{b}$ *has a unique solution.*
6. *For any n-column vector* $\mathbf{b}$*, the system* $\mathbf{Ax} = \mathbf{b}$ *has a solution.*

Proof. We prove (3)⇔(4) and leave the rest as an exercise. Multiply elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_k$ to $\mathbf{A}$ so that

$$\mathbf{R} = \mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_1 \mathbf{A}$$

is in reduced row echelon form. Then by Lemma 2.4.13, we have

$$\det(\mathbf{R}) = \det(\mathbf{E}_k) \det(\mathbf{E}_{k-1}) \cdots \det(\mathbf{E}_1) \det(\mathbf{A}).$$

Since determinant of elementary matrices are always nonzero, we have $\det(\mathbf{A})$ is nonzero if and only if $\det(\mathbf{R})$ is nonzero. It is easy to see that the determinant of a reduced row echelon matrix is nonzero if and only if it is the identity matrix $\mathbf{I}$. □

Theorem 2.4.16. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $n \times n$ *matrices. Then*

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B}).$$

Proof. If $\mathbf{A}$ is not invertible, then $\mathbf{AB}$ is not invertible and $\det(\mathbf{AB}) = 0 = \det(\mathbf{A})\det(\mathbf{B})$.
If $\mathbf{A}$ is invertible, then $\mathbf{A}$ is row equivalent to $\mathbf{I}$ and there exists elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_k$ such that

$$\mathbf{E}_k\mathbf{E}_{k-1}\cdots\mathbf{E}_1 = \mathbf{A}.$$

Hence

$$\begin{aligned}
\det(\mathbf{AB}) &= \det(\mathbf{E}_k\mathbf{E}_{k-1}\cdots\mathbf{E}_1\mathbf{B}) \\
&= \det(\mathbf{E}_k)\det(\mathbf{E}_{k-1})\cdots\det(\mathbf{E}_1)\det(\mathbf{B}) \\
&= \det(\mathbf{E}_k\mathbf{E}_{k-1}\cdots\mathbf{E}_1)\det(\mathbf{B}) \\
&= \det(\mathbf{A})\det(\mathbf{B}).
\end{aligned}$$

□

Definition 2.4.17 (Adjoint matrix)**.** *Let* $\mathbf{A}$ *be a square matrix. The* **adjoint matrix** *of* $\mathbf{A}$ *is*

$$\mathrm{adj}\mathbf{A} = [A_{ij}]^T,$$

where A_{ij} *is the* ij*-th cofactor of* $\mathbf{A}$*. In other words,*

$$[\mathrm{adj}\mathbf{A}]_{ij} = A_{ji}.$$

Theorem 2.4.18. *Let* $\mathbf{A}$ *be a square matrix. Then*

$$\mathbf{A}(\mathrm{adj}\mathbf{A}) = (\mathrm{adj}\mathbf{A})\mathbf{A} = \det(\mathbf{A})\mathbf{I},$$

where $\mathrm{adj}\mathbf{A}$ *is the adjoint matrix of* $\mathbf{A}$*. In particular if* $\mathbf{A}$ *is invertible, then*

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathrm{adj}\mathbf{A}.$$

Proof. The second statement follows easily from the first. For the first statement, we have

$$\begin{aligned}
[\mathbf{A}\mathrm{adj}\mathbf{A}]_{ij} &= \sum_{l=1}^{n} a_{il}[\mathrm{adj}\mathbf{A}]_{lj} \\
&= \sum_{l=1}^{n} a_{il}A_{jl} \\
&= \delta_{ij}\det(\mathbf{A})
\end{aligned}$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

Therefore $\mathbf{A}(\mathrm{adj}\mathbf{A}) = \det(A)\mathbf{I}$ and similarly $(\mathrm{adj}\mathbf{A})\mathbf{A} = \det(A)\mathbf{I}$. □

Example 2.4.19. *Find the inverse of*

$$\mathbf{A} = \begin{pmatrix} 4 & 3 & 2 \\ 5 & 6 & 3 \\ 3 & 5 & 2 \end{pmatrix}$$

Solution:

$$\det(\mathbf{A}) = 4\begin{vmatrix} 6 & 3 \\ 5 & 2 \end{vmatrix} - 3\begin{vmatrix} 5 & 3 \\ 3 & 2 \end{vmatrix} + 2\begin{vmatrix} 5 & 6 \\ 3 & 5 \end{vmatrix} = 4(-3) - 3(1) + 2(7) = -1,$$

$$\text{adj}\mathbf{A} = \begin{pmatrix} \begin{vmatrix} 6 & 3 \\ 5 & 2 \end{vmatrix} & -\begin{vmatrix} 3 & 2 \\ 5 & 2 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ 6 & 3 \end{vmatrix} \\ -\begin{vmatrix} 5 & 3 \\ 3 & 2 \end{vmatrix} & \begin{vmatrix} 4 & 2 \\ 3 & 2 \end{vmatrix} & -\begin{vmatrix} 4 & 2 \\ 5 & 3 \end{vmatrix} \\ \begin{vmatrix} 5 & 6 \\ 3 & 5 \end{vmatrix} & -\begin{vmatrix} 4 & 3 \\ 3 & 5 \end{vmatrix} & \begin{vmatrix} 4 & 3 \\ 5 & 6 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -3 & 4 & -3 \\ -1 & 2 & -2 \\ 7 & -11 & 9 \end{pmatrix}.$$

Therefore

$$\mathbf{A}^{-1} = \frac{1}{-1}\begin{pmatrix} -3 & 4 & -3 \\ -1 & 2 & -2 \\ 7 & -11 & 9 \end{pmatrix} = \begin{pmatrix} 3 & -4 & 3 \\ 1 & -2 & 2 \\ -7 & 11 & -9 \end{pmatrix}.$$

□

Theorem 2.4.20 (Cramer's rule)**.** *Consider the $n \times n$ linear system $\mathbf{Ax} = \mathbf{b}$, with*

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

where $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ are the column vectors of $\mathbf{A}$. If $\det(\mathbf{A}) \neq 0$, then the i-th entry of the unique solution $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is

$$x_i = \det(\mathbf{A})^{-1}\det(\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \end{bmatrix}),$$

where the matrix in the last factor is obtained by replacing the i-th column of $\mathbf{A}$ by $\mathbf{b}$.

Proof. For $i = 1, 2, \cdots, n$, we have

$$\begin{aligned} x_i &= [\mathbf{A}^{-1}\mathbf{b}]_i \\ &= \frac{1}{\det(\mathbf{A})}[(\text{adj}\mathbf{A})\mathbf{b}]_i \\ &= \frac{1}{\det(\mathbf{A})}\sum_{l=1}^{n} A_{li}b_l \\ &= \frac{1}{\det(\mathbf{A})}\det(\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \end{bmatrix}) \end{aligned}$$

□

Example 2.4.21. *Use Cramer's rule to solve the linear system*

$$\begin{cases} x_1 + 4x_2 + 5x_3 = 2 \\ 4x_1 + 2x_2 + 5x_3 = 3 \\ -3x_1 + 3x_2 - x_3 = 1 \end{cases}.$$

Solution:

$$\det(\mathbf{A}) = \begin{vmatrix} 1 & 4 & 5 \\ 4 & 2 & 5 \\ -3 & 3 & -1 \end{vmatrix} = 29.$$

Thus by Cramer's rule,

$$\begin{aligned}
x_1 &= \frac{1}{29}\begin{vmatrix} 2 & 4 & 5 \\ 3 & 2 & 5 \\ 1 & 3 & -1 \end{vmatrix} = \frac{33}{29} \\
x_2 &= \frac{1}{29}\begin{vmatrix} 1 & 2 & 5 \\ 4 & 3 & 5 \\ -3 & 1 & -1 \end{vmatrix} = \frac{35}{29} \\
x_3 &= \frac{1}{29}\begin{vmatrix} 1 & 4 & 2 \\ 4 & 2 & 3 \\ -3 & 3 & 1 \end{vmatrix} = -\frac{23}{29}
\end{aligned}$$

□

Exercise 2.4

1. Evaluate the following determinants.

   (a) $\begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix}$

   (b) $\begin{vmatrix} 1 & 4 & -2 \\ 3 & 2 & 0 \\ -1 & 4 & 3 \end{vmatrix}$

   (c) $\begin{vmatrix} 5 & 3 & 0 & 6 \\ 4 & 6 & 4 & 12 \\ 0 & 2 & -3 & 4 \\ 0 & 1 & -2 & 2 \end{vmatrix}$

   (d) $\begin{vmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \\ -2 & -2 & 3 & 3 \\ 1 & 2 & -2 & -3 \end{vmatrix}$

2. Suppose $\mathbf{A}$ is a $n \times n$ non-singular matrix and $\det(\mathbf{A}) = a$. Find the determinant of the following matrices in terms of n and a.

   (a) $\mathbf{A}^T$

   (b) $\mathbf{A}^2$

   (c) $-\mathbf{A}$

   (d) $3\mathbf{A}$

   (e) $\mathbf{A}^{-1}$

   (f) $\text{adj}\mathbf{A}$

3. For the given matrix $\mathbf{A}$, evaluate $\mathbf{A}^{-1}$ by finding the adjoint matrix $\text{adj}\mathbf{A}$ of $\mathbf{A}$.

   (a) $\mathbf{A} = \begin{pmatrix} 2 & 5 & 5 \\ -1 & -1 & 0 \\ 2 & 4 & 3 \end{pmatrix}$ (b) $\mathbf{A} = \begin{pmatrix} 2 & -3 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 2 \end{pmatrix}$ (c) $\mathbf{A} = \begin{pmatrix} 1 & 3 & 0 \\ -2 & -3 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

4. Use Cramer's rule to solve the following linear systems.

   (a) $\begin{cases} 4x_1 - x_2 - x_3 = 1 \\ 2x_1 + 2x_2 + 3x_3 = 10 \\ 5x_1 - 2x_2 - 2x_3 = -1 \end{cases}$

   (b) $\begin{cases} -x_1 + 2x_2 - 3x_3 = 1 \\ 2x_1 + x_3 = 0 \\ 3x_1 - 4x_2 + 4x_3 = 2 \end{cases}$

   (c) $\begin{cases} x_1 + 2x_3 = 6 \\ -3x_1 + 4x_2 + 6x_3 = 30 \\ -x_1 - 2x_2 + 3x_3 = 8 \end{cases}$

   (d) $\begin{cases} x_1 - 4x_2 + x_3 = 6 \\ 4x_1 - x_2 + 2x_3 = -1 \\ 2x_1 + 2x_2 - 3x_3 = -20 \end{cases}$

5. Show that

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-a)(c-b)$$

6. Let $a(t)$, $b(t)$, $c(t)$, $d(t)$ be differentiable functions of t. Prove that

$$\frac{d}{dt}\begin{vmatrix} a(t) & b(t) \\ c(t) & d(t) \end{vmatrix} = \begin{vmatrix} a'(t) & b'(t) \\ c(t) & d(t) \end{vmatrix} + \begin{vmatrix} a(t) & b(t) \\ c'(t) & d'(t) \end{vmatrix}$$

2.5 Linear equations and curve fitting

Given $n+1$ points on the coordinates plane with distinct x-coordinates, it is known that there exists a unique polynomial of degree at most n which fits the $n+1$ points. The formula for this polynomial which is called the **interpolation formula**, can be written in terms of determinant.

Theorem 2.5.1. *Let n be a non-negative integer, and $(x_0, y_0), (x_1, y_1), \cdots, (x_n, y_n)$ be $n+1$ points in $\mathbb{R}^2$ such that $x_i \neq x_j$ for any $i \neq j$. Then there exists unique polynomial*

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

of degree at most n such that $p(x_i) = y_i$ for all $0 \leq i \leq n$. The coefficients of $p(x)$ satisfy the linear system

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Moreover, we can write down the polynomial function $y = p(x)$ directly as

$$\begin{vmatrix} 1 & x & x^2 & \cdots & x^n & y \\ 1 & x_0 & x_0^2 & \cdots & x_0^n & y_0 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n & y_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n & y_n \end{vmatrix} = 0.$$

Proof. Expanding the determinant, one sees that the equation is of the form $y = p(x)$ where $p(x)$ is a polynomial of degree at most n. Observe when $(x, y) = (x_i, y_i)$ for some $0 \leq i \leq n$, two rows of the determinant would be the same and thus the determinant must be equal to zero. Moreover, it is well known that such polynomial is unique. □

Example 2.5.2. *Find the equation of straight line passes through the points (x_0, y_0) and (x_1, y_1).*

Solution: The equation of the required straight line is

$$\begin{aligned} \begin{vmatrix} 1 & x & y \\ 1 & x_0 & y_0 \\ 1 & x_1 & y_1 \end{vmatrix} &= 0 \\ (y_0 - y_1)x + (x_1 - x_0)y + (x_0y_1 - x_1y_0) &= 0 \end{aligned}$$

□

Example 2.5.3. *Find the cubic polynomial that interpolates the data points $(-1, 4), (1, 2), (2, 1)$ and $(3, 16)$.*

Solution: The required equation is

$$\begin{aligned}
\begin{vmatrix} 1 & x & x^2 & x^3 & y \\ 1 & -1 & 1 & -1 & 4 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 4 & 8 & 1 \\ 1 & 3 & 9 & 27 & 16 \end{vmatrix} &= 0 \\
\begin{vmatrix} 1 & x & x^2 & x^3 & y \\ 1 & -1 & 1 & -1 & 4 \\ 0 & 2 & 0 & 2 & -2 \\ 0 & 3 & 3 & 9 & -3 \\ 0 & 4 & 8 & 28 & 12 \end{vmatrix} &= 0 \\
&\vdots \\
\begin{vmatrix} 1 & x & x^2 & x^3 & y \\ 1 & 0 & 0 & 0 & 7 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 1 & 2 \end{vmatrix} &= 0 \\
-7 + 3x + 4x^2 - 2x^3 + y &= 0 \\
y &= 7 - 3x - 4x^2 + 2x^3
\end{aligned}$$

□

Using the same method, we can write down the equation of the circle passing through 3 given distinct non-colinear points directly without solving equations.

Example 2.5.4. *Find the equation of the circle that is determined by the points* $(-1, 5), (5, -3)$ *and* $(6, 4)$.

Solution: The equation of the required circle is

$$\begin{aligned}
\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ (-1)^2 + 5^2 & -1 & 5 & 1 \\ 5^2 + (-3)^2 & 5 & -3 & 1 \\ 6^2 + 4^2 & 6 & 4 & 1 \end{vmatrix} &= 0 \\
\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ 26 & -1 & 5 & 1 \\ 34 & 5 & -3 & 1 \\ 52 & 6 & 4 & 1 \end{vmatrix} &= 0 \\
&\vdots \\
\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ 20 & 0 & 0 & 1 \\ 4 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \end{vmatrix} &= 0 \\
x^2 + y^2 - 4x - 2y - 20 &= 0
\end{aligned}$$

□

Exercise 2.5

1. Find the equation of the parabola of the form $y = ax^2 + bx + c$ passing through the given set of three points.

(a) $(0,-5),(2,-1),(3,4)$ (b) $(-2,9),(1,3),(2,5)$ (c) $(-2,5),(-1,2),(1,-1)$

2. Find the equation of the circle passing through the given set of three points.

   (a) $(-1,-1),(6,6),(7,5)$ (b) $(3,-4),(5,10),(-9,12)$ (c) $(1,0),(0,-5),(-5,-4)$

3. Find the equation of a polynomial curve of degree 3 that passing through the points $(-1,3),(0,5),(1,7),(2,3)$.

4. Find the equation of a curve of the given form that passing through the given set of points.

   (a) $y = a + \frac{b}{x}$; $(1,5),(2,4)$

   (b) $y = ax + \frac{b}{x} + \frac{c}{x^2}$; $(1,2),(2,20),(4,41)$

   (c) $y = \frac{a}{x+b}$; $(1,2),(4,1)$

   (d) $y = \frac{ax+b}{cx+d}$; $(0,2),(1,1),(3,5)$

3 Vector spaces

3.1 Definition and examples

Consider the Euclidean space $\mathbb{R}^n$, the set of polynomials and the set of continuous functions on $[0,1]$. These sets look very differently. However, they share the same properties that similarly algebraic operations, namely addition and scalar multiplication, are defined on them. In mathematics, we call a set with these two algebraic structures a vector space.

Definition 3.1.1 (Vector space)**.** *A* **vector space** *over* $\mathbb{R}$ *consists of a set* V *and two algebraic operations addition and scalar multiplication such that*

1. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$*, for any* $\mathbf{u},\mathbf{v}\in V$
2. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$*, for any* $\mathbf{u},\mathbf{v},\mathbf{w}\in V$
3. *There exists* $\mathbf{0}\in V$ *such that* $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$*, for any* $\mathbf{u}\in V$
4. *For any* $\mathbf{u}\in V$*, there exists* $-\mathbf{u}\in V$ *such that* $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$
5. $a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v}$*, for any* $a\in\mathbb{R}$ *and* $\mathbf{u},\mathbf{v}\in V$
6. $(a+b)\mathbf{u}=a\mathbf{u}+b\mathbf{u}$*, for any* $a,b\in\mathbb{R}$ *and* $\mathbf{u}\in V$
7. $a(b\mathbf{u})=(ab)\mathbf{u}$*, for any* $a,b\in\mathbb{R}$ *and* $\mathbf{u}\in V$
8. $1\mathbf{u}=\mathbf{u}$*, for any* $\mathbf{u}\in V$

Example 3.1.2 (Euclidean space)**.** *The set*

$$\mathbb{R}^n=\left\{\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix} : x_i\in\mathbb{R}\right\}$$

with addition defined by

$$\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}+\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}=\begin{pmatrix} x_1+y_1\\ x_2+y_2\\ \vdots\\ x_n+y_n\end{pmatrix},$$

and scalar multiplication defined by

$$a\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix} ax_1\\ ax_2\\ \vdots\\ ax_n\end{pmatrix}$$

is a vector space which is called the **Euclidean space** *of dimension* n*.*

Example 3.1.3 (Matrix space)**.** *The set of all* $m\times n$ *matrices*

$$M_{m\times n}=\{\mathbf{A}:\mathbf{A} \textit{ is an } m\times n \textit{ matrix.}\}$$

with matrix addition and scaler multiplication is a vector space.

Example 3.1.4 (Space of polynomials over $\mathbb{R}$)**.** *The set of all polynomials*

$$P_n = \{a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} : a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}.\}$$

of degree less than n over $\mathbb{R}$ with addition

$$\begin{array}{rl} & (a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}) + (b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}) \\ = & (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_{n-1} + b_{n-1})x^{n-1} \end{array}$$

and scalar multiplication

$$a(a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}) = aa_0 + aa_1 x + \cdots + aa_{n-1}x^{n-1}$$

is a vector space.

Example 3.1.5 (Space of continuous functions)**.** *The set of all continuous functions*

$$C[a, b] = \{f : f \text{ is a continuous function on } [a, b]\}$$

on $[a, b]$ with addition and scalar multiplication defined by

$$\begin{array}{rcl} (f + g)(x) & = & f(x) + g(x) \\ (af)(x) & = & a(f(x)) \end{array}$$

is a vector space.

3.2 Subspaces

Definition 3.2.1. *Let W be a nonempty subset of the vector space V. Then W is a* **subspace** *of V if W itself is a vector space with the operations of addition and scalar multiplication defined in V.*

Theorem 3.2.2. *A nonempty subset W of a vector space V is a subspace of V if and only if it satisfies the following two conditions.*

1. *If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then $\mathbf{u} + \mathbf{v}$ is also in W.*
2. *If $\mathbf{u}$ is in W and c is a scalar, then $c\mathbf{u}$ is also in W.*

Example 3.2.3. *In the following examples, W is a vector subspace of V:*

1. *V is any vector space; $W = V$ or $\{\mathbf{0}\}$*
2. *$V = \mathbb{R}^n$; $W = \{(x_1, x_2, \cdots, x_n)^T \in V : a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0\}$, where $a_1, a_2, \cdots, a_n$ are fixed real numbers.*
3. *$V = M_{2\times 2}$; $W = \{\mathbf{A} = [a_{ij}] \in V : a_{11} + a_{22} = 0\}$.*
4. *V is the set $C[a, b]$ of all continuous functions on $[a, b]$; $W = \{f(x) \in V : f(a) = f(b) = 0\}$.*
5. *V is the set P_n of all polynomials of degree less than n; $W = \{p(x) \in V : p(0) = 0\}$.*
6. *V is the set P_n of all polynomials of degree less than n; $W = \{p(x) \in V : p'(0) = 0\}$.*

Example 3.2.4. *In the following examples, W is not a vector subspace of V:*

1. *$V = \mathbb{R}^2$; $W = \{(x_1, x_2)^T \in V : x_1 = 1\}$*
2. *$V = \mathbb{R}^n$; $W = \{(x_1, x_2, \cdots, x_n)^T \in V : x_1x_2 = 0\}$*
3. *$V = M_{2\times2}$; $W = \{\mathbf{A} \in V : \det(\mathbf{A}) = 0\}$*

Example 3.2.5. *Let $\mathbf{A} \in M_{m\times n}$, then the solution set of the homogeneous linear system*

$$\mathbf{Ax} = \mathbf{0}$$

is a subspace of $\mathbb{R}^n$. This subspace is called the **solution space** *of the system.*

Proposition 3.2.6. *Let U and W be two subspaces of a vector space V, then*

1. *$U \cap W = \{\mathbf{x} \in V : \mathbf{x} \in U \text{ and } \mathbf{x} \in W\}$ is subspace of V.*
2. *$U + W = \{\mathbf{u} + \mathbf{w} \in V : \mathbf{u} \in U \text{ and } \mathbf{w} \in W\}$ is subspace of V.*
3. *$U \cup W = \{\mathbf{x} \in V : \mathbf{x} \in U \text{ or } \mathbf{x} \in W\}$ is a subspace of V if and only if $U \subset W$ or $W \subset U$.*

Exercise 3.2

1. Determine whether the given subset W of $\mathbb{R}^3$ is a vector subspace of $\mathbb{R}^3$.
   (a) $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_2 = 0\}$
   (b) $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 1\}$
   (c) $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1x_2 = 0\}$
   (d) $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = 2x_2\}$
   (e) $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = x_3\}$
   (f) $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_2 + x_3 = 1\}$

2. Determine whether the given subset W of the set P_3 of polynomials of degree less than 3 is a vector subspace of P_3.
   (a) $W = \{p(x) \in P_3 : p(0) = 0\}$
   (b) $W = \{p(x) \in P_3 : p(1) = 0\}$
   (c) $W = \{p(x) \in P_3 : p(1) = 1\}$
   (d) $W = \{p(x) \in P_3 : p'(0) = 0\}$
   (e) $W = \{p(x) \in P_3 : p'(0) = p(0)\}$
   (f) $W = \{p(x) \in P_3 : p(0) = 2p(1)\}$

3. Determine whether the given subset W of the set $M_{2\times2}$ of 3×3 matrics is a vector subspace of $M_{3\times3}$.
   (a) $W = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times2} : a + d = 0\right\}$
   (b) $W = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times2} : ad = 1\right\}$
   (c) $W = \{\mathbf{A} \in M_{3\times3} : \det(\mathbf{A}) = 1\}$
   (d) $W = \{\mathbf{A} \in M_{3\times3} : \mathbf{A}^T = -\mathbf{A}\}$

3.3 Linear independence of vectors

Definition 3.3.1. *Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k \in V$. A* **linear combination** *of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ is a vector in V of the form*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k, \quad c_1, c_2, \cdots, c_n \in \mathbb{R}.$$

The **span** *of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ is the set of all linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ and is denoted by $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$. If W is a subspace of V and $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\} = W$, then we say that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ is a* **spanning set** *of W or $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$* **span** *the subspace W.*

Theorem 3.3.2. *Let* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k \in V$*. Then*

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$$

is a subspace of V*.*

Example 3.3.3. *Let* $V = \mathbb{R}^3$*.*

1. *If* $\mathbf{v}_1 = (1,0,0)^T$ *and* $\mathbf{v}_2 = (0,1,0)^T$*, then* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{(\alpha, \beta, 0)^T : \alpha, \beta \in \mathbb{R}\}$.
2. *If* $\mathbf{v}_1 = (1,0,0)^T$, $\mathbf{v}_2 = (0,1,0)^T$ *and* $\mathbf{v}_3 = (0,0,1)^T$*, then* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = V$.
3. *If* $\mathbf{v}_1 = (2,0,1)^T$ *and* $\mathbf{v}_2 = (0,1,-3)^T$*, then* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{(2\alpha, \beta, \alpha - 3\beta)^T : \alpha, \beta \in \mathbb{R}\}$.
4. *If* $\mathbf{v}_1 = (1,-1,0)^T$, $\mathbf{v}_2 = (0,1,-1)^T$ *and* $\mathbf{v}_3 = (-1,0,1)^T$*, then* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(x_1, x_2, x_3)^T : x_1 + x_2 + x_3 = 0\}$.

Example 3.3.4. *Let* $V = P_3$ *be the set of all polynomial of degree less than 3.*

1. *If* $\mathbf{v}_1 = x$ *and* $\mathbf{v}_2 = x^2$*, then* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{p(x) \in V : p(0) = 0\}$.
2. *If* $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = 3x - 2$ *and* $\mathbf{v}_3 = 2x + 1$*, then* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = P_2$.
3. *If* $\mathbf{v}_1 = 1 - x^2$, $\mathbf{v}_2 = x + 2$ *and* $\mathbf{v}_3 = x^2$*, then* $1 = \mathbf{v}_1 + \mathbf{v}_3$, $x = -2\mathbf{v}_1 + \mathbf{v}_2 - 2\mathbf{v}_3$ *and* $x^2 = \mathbf{v}_3$. *Thus* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ *contains* $\text{span}\{1, x, x^2\} = P_3$. *Therefore* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = P_3$.

Example 3.3.5. *Let* $\mathbf{w} = (2,-6,3)^T \in \mathbb{R}^3$, $\mathbf{v}_1 = (1,-2,-1)^T$ *and* $\mathbf{v}_2 = (3,-5,4)^T$*. Determine whether* $\mathbf{w} \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$*.*

Solution: Write

$$c_1 \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ -5 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ -6 \\ 3 \end{pmatrix},$$

that is

$$\begin{pmatrix} 1 & 3 \\ -2 & -5 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -6 \\ 3 \end{pmatrix}.$$

The augmented matrix

$$\left(\begin{array}{cc|c} 1 & 3 & 2 \\ -2 & -5 & -6 \\ -1 & 4 & 3 \end{array}\right)$$

can be reduced by elementary row operations to row echelon form

$$\left(\begin{array}{cc|c} 1 & 3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 19 \end{array}\right).$$

Since the system is inconsistent, we conclude that $\mathbf{w}$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.□

Example 3.3.6. *Let* $\mathbf{w} = (-7,7,11)^T \in \mathbb{R}^3$, $\mathbf{v}_1 = (1,2,1)^T$, $\mathbf{v}_2 = (-4,-1,2)^T$ *and* $\mathbf{v}_3 = (-3,1,3)^T$*. Express* $\mathbf{w}$ *as a linear combination of* $\mathbf{v}_1$, $\mathbf{v}_2$ *and* $\mathbf{v}_3$*.*

Solution: Write

$$c_1\begin{pmatrix}1\\2\\1\end{pmatrix}+c_2\begin{pmatrix}-4\\-1\\2\end{pmatrix}+c_3\begin{pmatrix}-3\\1\\3\end{pmatrix}=\begin{pmatrix}-7\\7\\11\end{pmatrix},$$

that is

$$\begin{pmatrix}1&-4&-3\\2&-1&1\\1&2&3\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}-7\\7\\11\end{pmatrix}.$$

The augmented matrix

$$\left(\begin{array}{ccc|c}1&-4&-3&-7\\2&-1&1&7\\1&2&3&11\end{array}\right)$$

has reduced row echelon form

$$\left(\begin{array}{ccc|c}1&0&1&5\\0&1&1&3\\0&0&0&0\end{array}\right).$$

The system has more than one solution. For example we can write

$$\mathbf{w}=5\mathbf{v}_1+3\mathbf{v}_2,$$

or

$$\mathbf{w}=3\mathbf{v}_1+\mathbf{v}_2+2\mathbf{v}_3.$$

□

Example 3.3.7. *Let* $\mathbf{v}_1=(1,-1,0)^T$, $\mathbf{v}_2=(0,1,-1)^T$ *and* $\mathbf{v}_3=(-1,0,1)^T$. *Observe that*

1. *One of the vectors is a linear combination of the other. For example*

$$\mathbf{v}_3=-\mathbf{v}_1-\mathbf{v}_2.$$

2. *The space* $\operatorname{span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ *has a smaller spanning set. For example*

$$\operatorname{span}\{\mathbf{v}_1,\mathbf{v}_2\}=\operatorname{span}\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}.$$

3. *There exists numbers* $c_1,c_2,c_3\in\mathbb{R}$, *not all zero, such that* $c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3=\mathbf{0}$. *For example*

$$\mathbf{v}_1+\mathbf{v}_2+\mathbf{v}_3=\mathbf{0}.$$

Definition 3.3.8. *The vectors* $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ *in a vector space* V *are said be be* **linearly independent** *if the equation*

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_k\mathbf{v}_k=\mathbf{0}$$

has only the trivial solution $c_1=c_2=\cdots=c_k=0$. *The vectors* $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ *are said be be* **linearly dependent** *if they are not linearly independent.*

Theorem 3.3.9. *Let* V *be a vector space and* $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k\in V$. *Then the following statements are equivalent.*

1. *The vectors* $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ *are linearly independent.*
2. *None of the vectors is a linear combination of the other vectors.*

3. *There does not exist a smaller spanning set of* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$.
4. *Every vector in* $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ *can be expressed in only one way as a linear combination of* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$.

Example 3.3.10. *The standard unit vectors*

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, 0, \cdots, 0)^T \\ \mathbf{e}_2 &= (0, 1, 0, \cdots, 0)^T \\ &\vdots \\ \mathbf{e}_n &= (0, 0, 0, \cdots, 1)^T \end{aligned}$$

are linearly independent in $\mathbb{R}^n$.

Example 3.3.11. *Let* $\mathbf{v}_1 = (1, 2, 2, 1)^T$, $\mathbf{v}_2 = (2, 3, 4, 1)^T$, $\mathbf{v}_3 = (3, 8, 7, 5)^T$ *be vectors in* $\mathbb{R}^4$. *Write the equation* $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$ *as the system*

$$\left\{ \begin{array}{rcrcrcl} c_1 & + & 2c_2 & + & 3c_3 & = & 0 \\ 2c_1 & + & 3c_2 & + & 8c_3 & = & 0 \\ 2c_1 & + & 4c_2 & + & 7c_3 & = & 0 \\ c_1 & + & c_2 & + & 5c_3 & = & 0 \end{array} \right. .$$

The augmented matrix of the system reduces to the row echelon form

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Thus the only solution is $c_1 = c_2 = c_3 = 0$. *Therefore* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *are linearly independent.*

Example 3.3.12. *Let* $\mathbf{v}_1 = (2, 1, 3)^T$, $\mathbf{v}_2 = (5, -2, 4)^T$, $\mathbf{v}_3 = (3, 8, -6)^T$ *and* $\mathbf{v}_4 = (2, 7, -4)^T$ *be vectors in* $\mathbb{R}^3$. *Write the equation* $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + c_4\mathbf{v}_4 = \mathbf{0}$ *as the system*

$$\left\{ \begin{array}{rcrcrcrcl} c_1 & + & 5c_2 & + & 3c_3 & + & 2c_4 & = & 0 \\ c_1 & - & 2c_2 & + & 8c_3 & + & 7c_4 & = & 0 \\ 3c_1 & + & 4c_2 & - & 6c_3 & - & 4c_4 & = & 0 \end{array} \right. .$$

Since there are more unknowns than equations and the system is homogeneous, it has a nontrivial solution. Therefore $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ *are linearly dependent.*

Theorem 3.3.13.

1. *Two nonzero vectors* $\mathbf{v}_1, \mathbf{v}_2 \in V$ *are linearly dependent if and only if they are proportional, i.e., there exists* $c \in \mathbb{R}$ *such that* $\mathbf{v}_2 = c\mathbf{v}_1$.
2. *If one of the vectors of* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k \in V$ *is zero, then* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *are linearly dependent.*
3. *Let* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ *be* n *vectors in* $\mathbb{R}^n$ *and*

$$\mathbf{A} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$$

be the $n \times n$ *matrix having them as its column vectors. Then* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ *are linearly independent if and only if* $\det(\mathbf{A}) \neq 0$.

4. *Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ be k vectors in $\mathbb{R}^n$, with $k > n$, then $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly dependent.*

Proof.

1. Obvious.
2. We may assume $\mathbf{v}_1 = \mathbf{0}$. Then
$$1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_k = \mathbf{0}.$$
Therefore $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly dependent.
3.
$$\begin{aligned} & \text{The vectors } \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n \text{ are linearly independent} \\ \Leftrightarrow\ & \text{The system } \mathbf{Ax} = \mathbf{0} \text{ has only trivial solution.} \\ \Leftrightarrow\ & \mathbf{A} \text{ is nonsingular} \\ \Leftrightarrow\ & \det(\mathbf{A}) \neq 0. \end{aligned}$$
4. Since the system
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$
has more unknowns than the number of equations, it must have nontrivial solution for $c_1, c_2, \cdots, c_k$. Therefore $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly dependent.

□

Exercise 3.3

1. Determine whether the given set of vectors are linearly independent in $\mathbb{R}^3$.
   (a) $\mathbf{v}_1 = (0, 2, -1), \mathbf{v}_2 = (2, -1, 3)$
   (b) $\mathbf{v}_1 = (1, 0, 1), \mathbf{v}_2 = (-2, 0, -2)$
   (c) $\mathbf{v}_1 = (1, 0, 0), \mathbf{v}_2 = (1, 1, 0), \mathbf{v}_3 = (1, 1, 1)$
   (d) $\mathbf{v}_1 = (1, -1, 0), \mathbf{v}_2 = (0, 1, -1), \mathbf{v}_3 = (-1, 0, 1)$
   (e) $\mathbf{v}_1 = (3, -1, -2), \mathbf{v}_2 = (2, 0, -1), \mathbf{v}_3 = (1, -3, -2)$
   (f) $\mathbf{v}_1 = (1, -2, 2), \mathbf{v}_2 = (3, 0, 1), \mathbf{v}_3 = (1, -1, 2)$
2. Suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent vectors in $\mathbb{R}^3$. Determine whether the given set $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ of vectors are linearly independent.
   (a) $\mathbf{u}_1 = \mathbf{v}_1, \mathbf{u}_2 = 2\mathbf{v}_2, \mathbf{u}_3 = 3\mathbf{v}_3$
   (b) $\mathbf{u}_1 = \mathbf{v}_1, \mathbf{u}_2 = \mathbf{v}_1 + \mathbf{v}_2, \mathbf{u}_3 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$
   (c) $\mathbf{u}_1 = \mathbf{v}_1 + 2\mathbf{v}_2, \mathbf{u}_2 = 2\mathbf{v}_1 - 4\mathbf{v}_2, \mathbf{u}_3 = -\mathbf{v}_1 + 3\mathbf{v}_2$
   (d) $\mathbf{u}_1 = 2\mathbf{v}_1 - \mathbf{v}_2, \mathbf{u}_2 = 2\mathbf{v}_2 - \mathbf{v}_3, \mathbf{u}_3 = 2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3$
3. Prove that if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent vectors, then $\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_1 + \mathbf{v}_3$ are linearly independent vectors.
4. Prove that if a set of vectors contains the zero vector, then it is linearly independent.
5. Prove that if S, T are two sets of vectors with $S \subset T$ and T is linearly independent, then S is linearly independent.
6. Let V be a vector space and W be a vector subspace of V. Suppose $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent vectors in W and $\mathbf{v}$ be a vector in V which does not lie in W. Prove that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k, \mathbf{v}$ are linearly independent.

3.4 Bases and dimension for vector spaces

Definition 3.4.1. *A set S of vectors in a vector space V is called a* **basis** *for V if*

1. *S is linearly independent, and*
2. *S spans V.*

Example 3.4.2.

1. *The vectors*

$$\begin{aligned}\mathbf{e}_1 &= (1,0,0,\cdots,0)^T\\ \mathbf{e}_2 &= (0,1,0,\cdots,0)^T\\ &\vdots\\ \mathbf{e}_n &= (0,0,0,\cdots,1)^T\end{aligned}$$

constitute a basis for $\mathbb{R}^n$ and is called the **standard basis** *for $\mathbb{R}^n$.*

2. *The vectors $\mathbf{v}_1 = (1,1,1)^T$, $\mathbf{v}_2 = (0,1,1)^T$ and $\mathbf{v}_3 = (2,0,1)^T$ constitute a basis for $\mathbb{R}^3$.*

Theorem 3.4.3. *If $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$, then any collection of m vectors in V, with $m > n$, are linearly dependent.*

Proof. Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m \in V$, $m > n$. Then we can write

$$\begin{aligned}\mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1n}\mathbf{v}_n\\ \mathbf{u}_2 &= a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{2n}\mathbf{v}_n\\ &\vdots\\ \mathbf{u}_m &= a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + \cdots + a_{mn}\mathbf{v}_n.\end{aligned}$$

We have

$$\begin{aligned}c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m &= \sum_{i=1}^{m}\left(c_i\sum_{j=1}^{n}a_{ij}\mathbf{v}_j\right)\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{m}c_ia_{ij}\right)\mathbf{v}_j\end{aligned}$$

Consider the

$$\left\{\begin{array}{ccccccccc} a_{11}c_1 & + & a_{21}c_2 & + & \cdots & + & a_{m1}c_m & = & 0\\ a_{12}c_1 & + & a_{22}c_2 & + & \cdots & + & a_{m2}c_m & = & 0\\ \vdots & & \vdots & & \ddots & & \vdots & = & \vdots\\ a_{1n}c_1 & + & a_{2n}c_2 & + & \cdots & + & a_{mn}c_m & = & 0\end{array}\right.$$

where $c_1, c_2, \cdots, c_m$ are variables. Since the number of unknowns is more than the number of equations, there exists nontrivial solution for $c_1, c_2, \cdots, c_m$ and

$$\sum_{i=1}^{m}c_ia_{ij} = 0, \text{ for } j = 1, 2, \cdots, n.$$

This implies that $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = \mathbf{0}$ and therefore $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$ are linearly dependent. □

Theorem 3.4.4. *Any two finite bases for a vector space consist of the same number of vectors.*

Proof. Let $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be two bases for V. Since $V = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ are linearly independent, we have $m \leq n$ by Theorem 3.4.3. Similarly, we have $n \leq m$. □

The above theorem enables us to define the dimension of a vector space.

Definition 3.4.5. *The* **dimension** *of a vector space V is the number of vectors of a finite basis of V. We say that V is* **of dimension** *n (or V is an n-dimensional vector space) if V has a basis consisting of n vectors. We say that V is an* **infinite dimensional** *vector space if it does not have a finite basis.*

Example 3.4.6.

1. *The Euclidean space $\mathbb{R}^n$ is of dimension n.*
2. *The polynomials $1, x, x^2, \cdots, x^{n-1}$ constitute a basis of the set for the set P_n of polynomials of degree less than n. Thus P_n is of dimension n.*
3. *The set of all $m \times n$ matrices $M_{m\times n}$ is of dimension mn.*
4. *The set of all continuous functions $C[a, b]$ is an infinite dimensional vector space.*

Theorem 3.4.7. *Let V be an n-dimension vector space and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be a subset of V consists of n vectors. Then the following statements for S are equivalent.*

1. *S is a basis for V.*
2. *S spans V.*
3. *S is linearly independent.*

Proof. We need to prove that S is linearly independent if and only if $\text{span}(S) = V$.
Suppose S is linearly independent and $\text{span}(S) \neq V$. Then there exists $\mathbf{v} \in V$ such that $\mathbf{v} \notin \text{span}(S)$. Since $S \cup \{\mathbf{v}\}$ contains $n + 1$ vectors, it is linearly dependent by Theorem 3.4.3. Thus there exists $c_1, c_2, \cdots, c_n, c_{n+1}$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v} = \mathbf{0}.$$

Now $c_{n+1} = 0$ since $\mathbf{v} \notin \text{span}(S)$. This contradicts to the assumption that $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ are linearly independent.
Suppose $\text{span}(S) = V$ and S is linearly dependent. Then by Theorem 3.3.9, there exists a proper subset $S' \subset S$ consists of k vectors, $k < n$, such that $\text{span}(S') = V$. By Theorem 3.4.3, any set of more than k vectors are linearly dependent. This contradicts to that V is of dimension n. □

Theorem 3.4.8. *Let V be an n-dimension vector space and let S be a subset of V. Then*

1. *If S is linearly independent, then S is contained in a basis for V.*
2. *If S spans V, then S contains a basis for V.*

Proof. 1. Suppose S is linearly independent. If $\mathrm{span}(S) = V$, then S is a basis for V. If $\mathrm{span}(S) \neq V$, then there exists $\mathbf{v}_1 \in V$ such that $\mathbf{v}_1 \notin \mathrm{span}(S)$. Now $S \cup \{\mathbf{v}_1\}$ is linearly independent. Similarly if $\mathrm{span}(S \cup \{\mathbf{v}_1\}) \neq V$, there exists $\mathbf{v}_2 \in V$ such that $S \cup \{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent. This process may be continued until $S \cup \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ contains n vectors. Then $S \cup \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ constitutes a basis for V.

2. Suppose S spans V. If S is linearly independent, then S is a basis for V. If S is linearly dependent, then there exists $\mathbf{v}_1 \in S$ which is a linear combination of the remaining vectors in S. After removing $\mathbf{v}_1$ from S, the remaining vectors will still span V. This process may be continued until we obtain a set of linearly independent vectors consisting of n vectors which constitutes a basis for V.

□

Theorem 3.4.9. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix. The set of solutions to the system*

$$\mathbf{A}\mathbf{x} = \mathbf{0}$$

form a vector subspace of $\mathbb{R}^n$*. The dimension of the solution space equals to the number of free variables.*

Example 3.4.10. *Find a basis for the solution space of the system*

$$\left\{\begin{array}{rcrcrcrcrcl} 3x_1 & + & 6x_2 & - & x_3 & - & 5x_4 & + & 5x_5 & = & 0 \\ 2x_1 & + & 4x_2 & - & x_3 & - & 3x_4 & + & 2x_5 & = & 0 \\ 3x_1 & + & 6x_2 & - & 2x_3 & - & 4x_4 & + & x_5 & = & 0. \end{array}\right.$$

Solution: The coefficient matrix $\mathbf{A}$ reduces to the row echelon form

$$\begin{pmatrix} 1 & 2 & 0 & -2 & 3 \\ 0 & 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The leading variables are x_1, x_3. The free variables are x_2, x_4, x_5. The set $\{(-2, 1, 0, 0, 0)^T$, $(2, 0, 1, 1, 0)^T$, $(-3, 0, -4, 0, 1)^T\}$ constitutes a basis for the solution space of the system. □

Exercise 3.4

1. Find a basis for the plane in $\mathbb{R}^3$ with the given equation.

   (a) $x + 2y - 4z = 0$ (b) $z = 3x - y$ (c) $3x + y = 0$

2. Find a basis for the solution space of the given homogeneous linear system.

   (a) $\left\{\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & 3x_3 & = & 0 \\ 2x_1 & - & 3x_2 & - & x_3 & = & 0 \end{array}\right.$

   (b) $\left\{\begin{array}{rcrcrcl} x_1 & + & 3x_2 & + & 4x_3 & = & 0 \\ 3x_1 & + & 8x_2 & + & 7x_3 & = & 0 \end{array}\right.$

   (c) $\left\{\begin{array}{rcrcrcrcl} x_1 & - & 3x_2 & + & 2x_3 & - & 4x_4 & = & 0 \\ 2x_1 & - & 5x_2 & + & 7x_3 & - & 3x_4 & = & 0 \end{array}\right.$

   (d) $\left\{\begin{array}{rcrcrcrcl} x_1 & - & 3x_2 & - & 9x_3 & - & 5x_4 & = & 0 \\ 2x_1 & + & x_2 & - & 4x_3 & + & 11x_4 & = & 0 \\ x_1 & + & 3x_2 & + & 3x_3 & + & 13x_4 & = & 0 \end{array}\right.$

$$\text{(e)}\ \left\{\begin{array}{rcrcrcrcl} x_1 & + & 5x_2 & + & 13x_3 & + & 14x_4 & = & 0 \\ 2x_1 & + & 5x_2 & + & 11x_3 & + & 12x_4 & = & 0 \\ 2x_1 & + & 7x_2 & + & 17x_3 & + & 19x_4 & = & 0 \end{array}\right.$$

$$\text{(f)}\ \left\{\begin{array}{rcrcrcrcl} x_1 & - & 3x_2 & - & 10x_3 & + & 5x_4 & = & 0 \\ x_1 & + & 4x_2 & + & 11x_3 & - & 2x_4 & = & 0 \\ x_1 & + & 3x_2 & + & 8x_3 & - & x_4 & = & 0 \end{array}\right.$$

3.5 Row and column spaces

Definition 3.5.1. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix.*

1. *The* **null space** $\mathrm{Null}(\mathbf{A})$ *of* $\mathbf{A}$ *is the solution space to* $\mathbf{Ax} = \mathbf{0}$*. In other words,* $\mathrm{Null}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R} : \mathbf{Ax} = \mathbf{0}.\}$*.*
2. *The* **row space** $\mathrm{Row}(\mathbf{A})$ *of* $\mathbf{A}$ *is the vector subspace of* $\mathbb{R}^n$ *spanned by the* m *row vectors of* $\mathbf{A}$*.*
3. *The* **column space** $\mathrm{Col}(\mathbf{A})$ *of* $\mathbf{A}$ *is the vector subspace of* $\mathbb{R}^m$ *spanned by the* n *column vectors of* $\mathbf{A}$*.*

It is easy to write down a basis for each of the above spaces for row echelon form.

Theorem 3.5.2. *Let* $\mathbf{R}$ *be a row echelon form. Then*

1. *The set of vectors obtained by setting one free variable equal to 1 and other free variables to be zero constitutes a basis for* $\mathrm{Null}(\mathbf{R})$*.*
2. *The set of non-zero rows constitutes a basis for* $\mathrm{Row}(\mathbf{R})$*.*
3. *The set of columns associated with lead variables constitutes a basis for* $\mathrm{Col}(\mathbf{R})$

Example 3.5.3. *Let*

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Find a basis for $\mathrm{Null}(\mathbf{A})$*,* $\mathrm{Row}(\mathbf{A})$ *and* $\mathrm{Col}(\mathbf{A})$*.*

Solution:

1. The set $\{(3,1,0,0,0)^T, (-3,0,2,-7,1)^T\}$ constitutes a basis for $\mathrm{Null}(\mathbf{A})$.
2. The set $\{(1,-3,0,0,3), (0,0,1,0,-2), (0,0,0,1,7)\}$ constitutes a basis for $\mathrm{Row}(\mathbf{A})$.
3. The set $\{(1,0,0,0)^T, (0,1,0,0)^T, (0,0,1,0)^T\}$ constitutes a basis for $\mathrm{Col}(\mathbf{A})$.

□

To find bases for the null space, row space and column space of a general matrix, we may find a row echelon form of the matrix and use the following theorem.

Theorem 3.5.4. *Let* $\mathbf{R}$ *be the row echelon form of* $\mathbf{A}$*. Then*

1. $\text{Null}(\mathbf{A}) = \text{Null}(\mathbf{R})$.
2. $\text{Row}(\mathbf{A}) = \text{Row}(\mathbf{R})$.
3. *The column vectors of* $\mathbf{A}$ *associated with the column containing the leading entries of* $\mathbf{R}$ *constitute a basis for* $\text{Col}(\mathbf{A})$.

Example 3.5.5. *Find a basis for the null space* $\text{Null}(\mathbf{A})$, *a basis for the row space* $\text{Row}(\mathbf{A})$ *and a basis for the column space* $\text{Col}(\mathbf{A})$ *where*

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 3 & 2 & 1 \\ 2 & -4 & 8 & 3 & 10 \\ 3 & -6 & 10 & 6 & 5 \\ 2 & -4 & 7 & 4 & 4 \end{pmatrix}.$$

Solution: The reduced row echelon form of $\mathbf{A}$ is

$$\begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus

1. the set $\{(2,1,0,0,0)^T, (-3,0,-2,4,1)^T\}$ constitutes a basis for $\text{Null}(\mathbf{A})$.
2. the set $\{(1,-2,0,0,3), (0,0,1,0,2), (0,0,0,1,-4)\}$ constitutes a basis for $\text{Row}(\mathbf{A})$.
3. the 1st, 3rd and 4th columns contain leading entries. Therefore the set $\{(1,2,3,2)^T$, $(3,8,10,7)^T$, $(2,3,6,4)^T\}$ constitutes a basis for $\text{Col}(\mathbf{A})$.

□

Definition 3.5.6. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix. The dimension of*

1. *the solution space of* $\mathbf{Ax} = \mathbf{0}$ *is called the* **nullity** *of* $\mathbf{A}$.
2. *the row space is called the* **row rank** *of* $\mathbf{A}$.
3. *the column space is called the* **column rank** *of* $\mathbf{A}$.

To find the above three quantities of a matrix, we have the following theorem which is a direct consequence of Theorem 3.5.2 and Theorem 3.5.4.

Theorem 3.5.7. *Let* $\mathbf{A}$ *be a matrix.*

1. *The nullity of* $\mathbf{A}$ *is equal to the number of free variables.*
2. *The row rank of* $\mathbf{A}$ *is equal to the number of lead variables.*
3. *The column rank of* $\mathbf{A}$ *is equal to the number of lead variables.*

Now we can state two important theorems for general matrices.

Theorem 3.5.8. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix. Then the row rank of* $\mathbf{A}$ *is equal to the column rank of* $\mathbf{A}$.

Proof. Both of them are equal to the number of leading entries of the reduced row echelon form of $\mathbf{A}$. □

The common value of the row and column rank of the matrix $\mathbf{A}$ is called the **rank** of $\mathbf{A}$ and is denoted by rank($\mathbf{A}$). The nullity of $\mathbf{A}$ is denoted by nullity($\mathbf{A}$). The rank and nullity of a matrix is related in the following way.

Theorem 3.5.9 (Rank-Nullity Theorem)**.** *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix. Then*

$$\operatorname{rank}(\mathbf{A}) + \operatorname{nullity}(\mathbf{A}) = n$$

where rank($\mathbf{A}$) *and* nullity($\mathbf{A}$) *are the rank and nullity of* $\mathbf{A}$ *respectively.*

Proof. The nullity of $\mathbf{A}$ is equal to the number of free variables of the reduced row echelon form of $\mathbf{A}$. Now the left hand side is the sum of the number of leading variables and free variables and is of course equal to n. □

We end this section by proving a theorem which will be used in Section 5.4.

Theorem 3.5.10. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two matrices such that* $\mathbf{AB}$ *is defined. Then*

$$\operatorname{nullity}(\mathbf{B}) \leq \operatorname{nullity}(\mathbf{AB}) \leq \operatorname{nullity}(\mathbf{A}) + \operatorname{nullity}(\mathbf{B})$$

Proof. It is obvious that Null($\mathbf{B}$) $\subset$ Null($\mathbf{AB}$). Thus we have nullity($\mathbf{B}$) $\leq$ nullity($\mathbf{AB}$). Observe that Null($\mathbf{AB}$) $= \{\mathbf{v} : \mathbf{Bv} \in \operatorname{Null}(\mathbf{A})\}$. Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ be vectors such that $\{\mathbf{Bu}_1, \mathbf{Bu}_2, \cdots, \mathbf{Bu}_k\}$ is a basis for Null($\mathbf{A}$)$\cap$Col($\mathbf{B}$) and $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l\}$ be a basis for Null($\mathbf{B}$). We are going to prove that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ constitute a basis for Null($\mathbf{AB}$). First we prove that they are linearly independent. Suppose

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k + d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_l\mathbf{v}_l = \mathbf{0}$$

Multiplying $\mathbf{B}$ from the left, we have

$$c_1\mathbf{Bu}_1 + c_2\mathbf{Bu}_2 + \cdots + c_k\mathbf{Bu}_k + d_1\mathbf{Bv}_1 + d_2\mathbf{Bv}_2 + \cdots + d_l\mathbf{Bv}_l = \mathbf{0}$$

and since $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l \in$ Null($\mathbf{B}$), we obtain

$$c_1\mathbf{Bu}_1 + c_2\mathbf{Bu}_2 + \cdots + c_k\mathbf{Bu}_k = \mathbf{0}$$

This implies that $c_1 = c_2 = \cdots = c_k = 0$ since $\mathbf{Bu}_1, \mathbf{Bu}_2, \cdots, \mathbf{Bu}_k$ are linearly independent. Thus

$$d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_l\mathbf{v}_l = \mathbf{0}$$

and consequently $d_1 = d_2 = \cdots = d_l = 0$ since $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ are linearly independent. Hence $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ are linearly independent.

Second we prove that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ span Null($\mathbf{AB}$). For any $\mathbf{v} \in$ Null($\mathbf{AB}$), we have $\mathbf{Bv} \in$ Null($\mathbf{A}$) $\cap$ Col($\mathbf{B}$). Since $\mathbf{Bu}_1, \mathbf{Bu}_2, \cdots, \mathbf{Bu}_k$ span Null($\mathbf{A}$) $\cap$ Col($\mathbf{B}$), there exists $c_1, c_2, \cdots, c_k$ such that

$$\mathbf{Bv} = c_1\mathbf{Bu}_1 + c_2\mathbf{Bu}_2 + \cdots + c_k\mathbf{Bu}_k$$

It follows that

$$\mathbf{v} - (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) \in \operatorname{Null}(\mathbf{B})$$

and since $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ span Null($\mathbf{B}$), there exists $d_1, d_2, \cdots, d_l$ such that

$$\mathbf{v} - (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) = d_1\mathbf{v}_1 + \mathbf{v}_2 + \cdots + d_l\mathbf{v}_l$$

Thus

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k + d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_l\mathbf{v}_l$$

This implies that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ span Null($\mathbf{AB}$). Hence we completed the proof that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l$ constitute a basis for Null($\mathbf{AB}$).

Observe that $k = \dim(\mathrm{Null}(\mathbf{A}) \cap \mathrm{Col}(\mathbf{B})) \leq \mathrm{nullity}(\mathbf{A})$ and $l = \mathrm{nullity}(\mathbf{B})$. Therefore we have

$$\mathrm{nullity}(\mathbf{AB}) = k + l \leq \mathrm{nullity}(\mathbf{A}) + \mathrm{nullity}(\mathbf{B})$$

□

Exercise 3.5

1. Find a basis for the null space , a basis for the row space and a basis for the column space for the given matrices.

   (a) $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 5 & -9 \\ 2 & 5 & 2 \end{pmatrix}$

   (b) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 3 & 1 & -3 & 4 \\ 2 & 5 & 11 & 12 \end{pmatrix}$

   (c) $\begin{pmatrix} 3 & -6 & 1 & 3 & 4 \\ 1 & -2 & 0 & 1 & 2 \\ 1 & -2 & 2 & 0 & 3 \end{pmatrix}$

   (d) $\begin{pmatrix} 1 & 1 & -1 & 7 \\ 1 & 4 & 5 & 16 \\ 1 & 3 & 3 & 13 \\ 2 & 5 & 4 & 23 \end{pmatrix}$

   (e) $\begin{pmatrix} 1 & -2 & -3 & -5 \\ 1 & 4 & 9 & 2 \\ 1 & 3 & 7 & 1 \\ 2 & 2 & 6 & -3 \end{pmatrix}$

   (f) $\begin{pmatrix} 1 & 1 & 3 & 3 & 1 \\ 2 & 3 & 7 & 8 & 2 \\ 2 & 3 & 7 & 8 & 3 \\ 3 & 1 & 7 & 5 & 4 \end{pmatrix}$

   (g) $\begin{pmatrix} 1 & 1 & 3 & 0 & 3 \\ -1 & 0 & -2 & 1 & -1 \\ 2 & 3 & 7 & 1 & 8 \\ -2 & 4 & 0 & 7 & 6 \end{pmatrix}$

2. Find a basis for the subspace spanned by the given set of vectors.

   (a) $\mathbf{v}_1 = (1, 3, -2, 4), \mathbf{v}_2 = (2, -1, 3, 1), \mathbf{v}_3 = (5, 1, 4, 6)$

   (b) $\mathbf{v}_1 = (1, -1, 2, 3), \mathbf{v}_2 = (2, 3, 4, 1), \mathbf{v}_3 = (1, 1, 2, 1), \mathbf{v}_4 = (4, 1, 8, 7)$

   (c) $\mathbf{v}_1 = (3, 2, 2, 2), \mathbf{v}_2 = (2, 1, 2, 1), \mathbf{v}_3 = (4, 3, 2, 3), \mathbf{v}_4 = (1, 2, -2, 4)$

   (d) $\mathbf{v}_1 = (1, -2, 1, 1, 2), \mathbf{v}_2 = (-1, 3, 0, 2, -2), \mathbf{v}_3 = (0, 1, 1, 3, 4), \mathbf{v}_4 = (1, 2, 5, 13, 5)$

   (e) $\mathbf{v}_1 = (1, -3, 4, -2, 5), \mathbf{v}_2 = (2, -6, 9, -1, 8), \mathbf{v}_3 = (2, -6, 9, -1, 9), \mathbf{v}_4 = (-1, 3, -4, 2, -5)$

3. Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $n \times k$ matrix. Let $r_A = \mathrm{rank}(\mathbf{A})$, $r_B = \mathrm{rank}(\mathbf{B})$ and r_{AB} be the rank of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{AB}$ respectively. Prove that

   $$r_A + r_B - n \leq r_{AB} \leq \min(r_A, r_B)$$

   where $\min(r_A, r_B)$ denotes the minimum of r_A and r_B.

3.6 Orthogonal vectors in $\mathbb{R}^n$

Definition 3.6.1. *An* **inner product** *on a vector space V is a function that associates with each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ a scalar $\langle \mathbf{u}, \mathbf{v} \rangle$ such that, for any scalar c and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$,*

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ *and* $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$

An **inner product space** *is a vector space V together with an inner product defined on V.*

Example 3.6.2.

1. *(Dot product) Let $\mathbf{u} = (u_1, \cdots, u_n)^T$ and $\mathbf{v} = (v_1, \cdots, v_n)^T$ be vectors in $\mathbb{R}^n$. The* **dot product** *defined by*
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$
is an inner product on $\mathbb{R}^n$.

2. *Let P_n be the set of all polynomials of degree less than n. Let $x_1, x_2, \cdots, x_n$ be n distinct real numbers and $p(x), q(x) \in P_n$. Then*
$$\langle p(x), q(x) \rangle = p(x_1)q(x_1) + p(x_2)q(x_2) + \cdots + p(x_n)q(x_n)$$
defines an inner product on P_n.

3. *Let $C[a, b]$ be the set of all continuous function on $[a, b]$ and $f(x), g(x) \in C[a, b]$. Then*
$$\langle f(x), g(x) \rangle = \int_a^b f(x)g(x)dx$$
defines an inner product on $C[a, b]$.

Definition 3.6.3. *Let V be an inner product space and $\mathbf{u} \in V$. The* **length***, or* **magnitude***, of $\mathbf{u}$ is defined as*
$$|\mathbf{u}| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$
The **distance** *between two vectors $\mathbf{u}$ and $\mathbf{v}$ in V is defined as*
$$|\mathbf{u} - \mathbf{v}|.$$

Theorem 3.6.4 (Cauchy-Schwarz inequality)**.** *Let V be an inner product space and $\mathbf{u}, \mathbf{v} \in V$. Then*
$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq |\mathbf{u}||\mathbf{v}|$$
and equality holds if and only if $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = t\mathbf{u}$ for some scalar t.

Proof. When $\mathbf{u} = \mathbf{0}$, the inequality is satisfied trivially. If $\mathbf{u} \neq \mathbf{0}$, then for any real number t, we have

$$\begin{array}{rcl} |t\mathbf{u} - \mathbf{v}|^2 & \geq & 0 \\ \langle (t\mathbf{u} - \mathbf{v}), (t\mathbf{u} - \mathbf{v}) \rangle & \geq & 0 \\ t^2\langle \mathbf{u}, \mathbf{u} \rangle - 2t\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle & \geq & 0 \end{array}$$

Thus the quadratic function $t^2\langle \mathbf{u},\mathbf{u}\rangle - 2t\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle$ is always non-negative. Hence its discriminant

$$4(\langle \mathbf{u},\mathbf{v}\rangle)^2 - 4(\langle \mathbf{u},\mathbf{u}\rangle)(\langle \mathbf{v},\mathbf{v}\rangle)$$

is non-positive. Therefore

$$\begin{aligned} (\langle \mathbf{u},\mathbf{v}\rangle)^2 &\leq (\langle \mathbf{u},\mathbf{u}\rangle)(\langle \mathbf{v},\mathbf{v}\rangle) \\ |\langle \mathbf{u},\mathbf{v}\rangle|^2 &\leq |\mathbf{u}|^2|\mathbf{v}|^2 \\ |\langle \mathbf{u},\mathbf{v}\rangle| &\leq |\mathbf{u}||\mathbf{v}| \end{aligned}$$

and equality holds if and only if $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = t\mathbf{u}$ for some t. □

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in an inner product space, by Cauchy-Schwarz inequality we have

$$-1 \leq \frac{\langle \mathbf{u},\mathbf{v}\rangle}{|\mathbf{u}||\mathbf{v}|} \leq 1.$$

This enables us to make the following definition.

Definition 3.6.5. *Let* $\mathbf{u},\mathbf{v}$ *be nonzero vectors in an inner product space. Then the angle between* $\mathbf{u}$ *and* $\mathbf{v}$ *is the unique angle* θ *between* 0 *and* π *inclusively such that*

$$\cos\theta = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{|\mathbf{u}||\mathbf{v}|}.$$

Definition 3.6.6. *We say that two vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are* **orthogonal** *in an inner product space* V *if* $\langle \mathbf{u},\mathbf{v}\rangle = 0$.

Theorem 3.6.7 (Triangle inequality)**.** *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be vectors in an inner product space. Then*

$$|\mathbf{u}+\mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

Proof. We apply Cauchy-Schwarz inequality to find that

$$\begin{aligned} |\mathbf{u}+\mathbf{v}|^2 &= \langle \mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}\rangle \\ &= \langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \\ &\leq |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2 \\ &= (|\mathbf{u}| + |\mathbf{v}|)^2. \end{aligned}$$

□

Theorem 3.6.8. *Let* $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ *be mutually orthogonal nonzero vectors in an inner product space* V*. Then they are linearly independent.*

Proof. Suppose

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

For each $1 \leq i \leq k$, we take the inner product of each side with $\mathbf{v}_i$, we have

$$c_i\langle \mathbf{v}_i,\mathbf{v}_i\rangle = 0.$$

Since $\mathbf{v}_i$ is a nonzero vector, we have $c_i = 0$. Thus $c_1 = c_2 = \cdots = c_k = 0$ and therefore $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ are linearly independent. □

Definition 3.6.9 (Orthogonal complement of a subspace)**.** *Let W be a vector subspace of an inner product space V. We say that a vector $\mathbf{u} \in V$ is* **orthogonal** *to W if $\mathbf{u}$ is orthogonal to every vector in W. The* **orthogonal complement** *of W in V is*

$$W^{\perp} = \{\mathbf{u} \in V : \langle \mathbf{u}, \mathbf{w} \rangle = 0, \text{ for all } \mathbf{w} \in W\}.$$

Theorem 3.6.10 (Properties of orthogonal complements)**.** *Let W be a vector subspace of an inner product space V. Then*

1. *$W^{\perp}$ is a vector subspace of V.*
2. *$W^{\perp} \cap W = \{\mathbf{0}\}$*
3. *If V is finite dimensional, then $\dim(W) + \dim(W^{\perp}) = \dim(V)$.*
4. *$W \subset (W^{\perp})^{\perp}$. If V is finite dimensional, then $(W^{\perp})^{\perp} = W$.*
5. *If S spans W, then $\mathbf{u} \in W^{\perp}$ if and only if $\mathbf{u}$ is orthogonal to every vector in S.*

Theorem 3.6.11. *Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ be (column) vectors in $\mathbb{R}^n$ and $W = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$. Then*

$$W^{\perp} = \operatorname{Null}(\mathbf{A})$$

where $\mathbf{A}$ is the $m \times n$ matrix with row vectors $\mathbf{v}_1^T, \mathbf{v}_2^T, \cdots, \mathbf{v}_m^T$.

Proof. For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned} \mathbf{x} \in W^{\perp} &\Leftrightarrow \langle \mathbf{v}_i, \mathbf{x} \rangle = 0 \text{ for any } i = 1, 2, \cdots, m \\ &\Leftrightarrow \mathbf{A}\mathbf{x} = \mathbf{0} \\ &\Leftrightarrow \mathbf{x} \in \operatorname{Null}(\mathbf{A}). \end{aligned}$$

□

To find a basis for the orthogonal complement $W^{\perp}$ of a subspace of the form $W = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$, we may write down a matrix $\mathbf{A}$ using $\mathbf{v}_1^T, \mathbf{v}_2^T, \cdots, \mathbf{v}_m^T$ as row vectors and then find a basis for $\operatorname{Null}(\mathbf{A})$.

Example 3.6.12. *Let $W = \operatorname{span}\{(1, -3, 5)^T\}$. Find a basis for $W^{\perp}$.*

Solution: Using $(1, -3, 5)$ as row vector, we obtain $\mathbf{A} = (1, -3, 5)$ which is in reduced row echelon form. Thus the vectors $(3, 1, 0)^T$ and $(-5, 0, 1)^T$ constitute a basis for $W^{\perp} = \operatorname{null}(\mathbf{A}^T)$. □

Example 3.6.13. *Let W be the subspace spanned by $(1, 2, 1, -3, -3)^T$ and $(2, 5, 6, -10, -12)^T$. Find a basis for $W^{\perp}$.*

Solution: Using $(1, 2, 1, -3, -3)$ and $(2, 5, 6, -10, -12)$ as row vectors, we obtain

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & -3 & -3 \\ 2 & 5 & 6 & -10 & -12 \end{pmatrix}$$

which has reduced row echelon form

$$\begin{pmatrix} 1 & 0 & -7 & 5 & 9 \\ 0 & 1 & 4 & -4 & -6 \end{pmatrix}.$$

Thus the vectors $(7, -4, 1, 0, 0)^T, (-5, 4, 0, 1, 0)^T, (-9, 6, 0, 0, 1)^T$ constitute a basis for $W^{\perp}$. □

Example 3.6.14. *Find a basis for the orthogonal complement of the subspace spanned by* $(1,2,-1,1)^T$, $(2,4,-3,0)^T$ *and* $(1,2,1,5)^T$.

Solution: Using $(1,2,-1,1)$, $(2,4,-3,0)$, $(1,2,1,5)$ as row vectors, we obtain

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -3 & 0 \\ 1 & 2 & 1 & 5 \end{pmatrix}$$

which has reduced row echelon form

$$\begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus the vectors $(-2,1,0,0)^T$ and $(-3,0,-2,1)^T$ constitute a basis for

$$\text{span}\{(1,2,-1,1)^T, (2,4,-3,0)^T, (1,2,1,5)^T\}^\perp.$$

$\square$

Exercise 3.6

1. Find a basis for the orthogonal complement of the subspace of the Euclidean space spanned by given set of vectors.
   (a) $\{(1,2,3)\}$
   (b) $\{(1,-2,-3,5)\}$
   (c) $\{(1,3,2,4),(2,7,7,3)\}$
   (d) $\{(1,-3,3,5),(2,-5,9,3)\}$
   (e) $\{(1,2,5,2,3),(3,7,11,9,5)\}$
   (f) $\{(2,5,5,4,3),(3,7,8,8,8)\}$
   (g) $\{(1,2,3,1,3),(1,3,4,3,6),(2,2,4,3,5)\}$
   (h) $\{(1,1,1,1,3),(2,3,1,4,7),(5,3,7,1,5)\}$

2. Prove that for any vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space V, we have
   (a) $|\mathbf{u}+\mathbf{v}|^2 + |\mathbf{u}-\mathbf{v}|^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2$
   (b) $|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2 = 4\langle\mathbf{u},\mathbf{v}\rangle$

3. Let V be an inner product space. Prove that for any vector subspace $W \subset V$, we have $W \cap W^\perp = \{\mathbf{0}\}$.

4. Let V be an inner product space. Prove that for any vector subspace $W \subset V$, we have $W \subset (W^\perp)^\perp$. (Note that in general $(W^\perp)^\perp \neq W$.)

5. Let V be a vector space and $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k \in V$ be non-zero vectors in V such that $\langle\mathbf{v}_i, \mathbf{v}_i\rangle = 0$ for any $i, j = 1, 2, \cdots, k$ with $i \neq j$. Prove that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent.

4 Second and higher order linear equations

4.1 Second order linear equations

In first part of this chapter, we consider second order linear ordinary linear equations, i.e., a differential equation of the form

$$\frac{d^2y}{dt^2} + p(t)\frac{dy}{dt} + q(t)y = g(t)$$

where $p(t)$ and $q(t)$ are continuous functions. We may let

$$L[y] = y'' + p(t)y' + q(t)y$$

and write the equation as the form

$$L[y] = g(t).$$

The above equation is said to be **homogeneous** if $g(t) = 0$ and the equation

$$L[y] = 0$$

is called the **associated homogeneous equation**. First we state two fundamental results of second order linear ODE.

Theorem 4.1.1 (Existence and uniqueness of solution)**.** *Let I be an open interval and $t_o \in I$. Let $p(t), q(t)$ and $g(t)$ be continuous functions on I. Then for any real numbers y_0 and y_0', the initial value problem*

$$\begin{cases} y'' + p(t)y' + q(t)y = g(t), & t \in I \\ y(t_0) = y_0, \quad y'(t_0) = y_0' \end{cases}$$

has a unique solution on I.

The proof of the above theorem needs some hard analysis and is omitted. But the proof of the following theorem is simple and is left to the readers.

Theorem 4.1.2 (Principle of superposition)**.** *If y_1 and y_2 are two solutions to the homogeneous equation*

$$L[y] = 0,$$

then $y = c_1y_1 + c_2y_2$ is also a solution for any constants c_1 and c_2.

The principle of superposition implies that the solutions of a homogeneous equation form a vector space. This suggests us finding a basis for the solution space. Let's recall the definition of linear independency for functions (See Definition 3.3.8 for linear independency for vectors in a general vector space).

Definition 4.1.3. *Two functions $u(t)$ and $v(t)$ are said to be* **linearly dependent** *if there exists constants c_1 and c_2, not both zero, such that $c_1u(t) + c_2v(t) = 0$ for all $t \in I$. They are said to be* **linearly independent** *if they are not linearly dependent.*

Definition 4.1.4 (Fundamental set of solutions)**.** *We say that two solutions y_1 and y_2 form a* **fundamental set of solutions** *of a second order homogeneous linear differential equation if they are linearly independent.*

Definition 4.1.5 (Wronskian)**.** *Let y_1 and y_2 be two differentiable functions. Then we define the* **Wronskian** *(or* **Wronskian determinant***) of y_1, y_2 to be the function*

$$W(t) = W(y_1, y_2)(t) = \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix} = y_1(t)y_2'(t) - y_1'(t)y_2(t).$$

Wronskian is used to determine whether a pair of solutions is linearly independent.

Theorem 4.1.6. *Let $u(t)$ and $v(t)$ be two differentiable functions on open interval I. If $W(u,v)(t_0) \neq 0$ for some $t_0 \in I$, then u and v are linearly independent.*

Proof. Suppose $c_1u(t) + c_2v(t) = 0$ for all $t \in I$ where c_1, c_2 are constants. Then we have

$$\begin{cases} c_1u(t_0) + c_2v(t_0) & = & 0, \\ c_1u'(t_0) + c_2v'(t_0) & = & 0. \end{cases}$$

In other words,

$$\begin{pmatrix} u(t_0) & v(t_0) \\ u'(t_0) & v'(t_0) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Now the matrix

$$\begin{pmatrix} u(t_0) & v(t_0) \\ u'(t_0) & v'(t_0 \end{pmatrix}$$

is non-singular since its determinant $W(u,v)(t_0)$ is non-zero by the assumption. This implies by Theorem 2.4.8 that $c_1 = c_2 = 0$. Therefore $u(t)$ and $v(t)$ are linearly independent. □

Remark: The converse of the above theorem is false. For example take $u(t) = t^3$, $v(t) = |t|^3$. Then $W(u,v)(t) = 0$ for any $t \in \mathbb{R}$ but $u(t)$, $v(t)$ are not linearly independent.

Example 4.1.7. *The functions $y_1(t) = e^t$ and $y_2(t) = e^{-2t}$ form a fundamental set of solutions of*

$$y'' + y' - 2y = 0$$

since $W(y_1, y_2) = e^t(-2e^{-2t}) - e^t(e^{-2t}) = -3e^{-t}$ is not identically zero.

Example 4.1.8. *The functions $y_1(t) = e^t$ and $y_2(t) = te^t$ form a fundamental set of solutions of*

$$y'' - 2y' + y = 0$$

since $W(y_1, y_2) = e^t(te^t + e^t) - e^t(te^t) = e^{2t}$ is not identically zero.

Example 4.1.9. *The functions $y_1(t) = 3$, $y_2(t) = \cos^2 t$ and $y_3(t) = -2\sin^2 t$ are linearly dependent since*

$$2(3) + (-6)\cos^2 t + 3(-2\sin^2 t) = 0.$$

One may justify that the Wronskian

$$W(y_1, y_2, y_3) = \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix} = 0.$$

Example 4.1.10. *Show that $y_1(t) = t^{1/2}$ and $y_2(t) = t^{-1}$ form a fundamental set of solutions of*

$$2t^2y'' + 3ty' - y = 0, \quad t > 0.$$

Solution: It is easy to check that y_1 and y_2 are solutions to the equation. Now

$$W(y_1, y_2)(t) = \begin{vmatrix} t^{1/2} & t^{-1} \\ \frac{1}{2}t^{-1/2} & -t^{-2} \end{vmatrix} = -\frac{3}{2}t^{-3/2}$$

is not identically zero. We conclude that y_1 and y_2 form a fundamental set of solutions of the equation. □

Theorem 4.1.11 (Abel's Theorem). *If y_1 and y_2 are solutions to the second order homogeneous equation*

$$L[y] = y'' + p(t)y' + q(t)y = 0,$$

where p and q are continuous on an open interval I, then

$$W(y_1, y_2)(t) = c\exp\left(-\int p(t)dt\right),$$

where c is a constant that depends on y_1 and y_2. Furthermore, $W(y_1, y_2)(t)$ is either identically zero on I or never zero on I.

Proof. Since y_1 and y_2 are solutions, we have

$$\begin{cases} y_1'' + p(t)y_1' + q(t)y_1 &= 0 \\ y_2'' + p(t)y_2' + q(t)y_2 &= 0. \end{cases}$$

If we multiply the first equation by $-y_2$, multiply the second equation by y_1 and add the resulting equations, we get

$$\begin{aligned} (y_1y_2'' - y_1''y_2) + p(t)(y_1y_2' - y_1'y_2) &= 0 \\ W' + p(t)W &= 0 \end{aligned}$$

which is a first-order linear and separable differential equation with solution

$$W(t) = c\exp\left(-\int p(t)dt\right),$$

where c is a constant. Since the value of the exponential function is never zero, $W(y_1, y_2)(t)$ is either identically zero on I (when $c = 0$) or never zero on I (when $c \neq 0$). □

Let y_1 and y_2 be two differentiable functions. In general, we cannot conclude that their Wronskian $W(t)$ is not identically zero purely from their linear independency. However, if y_1 and y_2 are solutions to a second order homogeneous linear differential equation, then $W(t_0) \neq 0$ for some $t_0 \in I$ provided y_1 and y_2 are linearly independent.

Theorem 4.1.12. *Suppose y_1 and y_2 are solutions to the second order homogeneous equation*

$$L[y] = y'' + p(t)y' + q(t)y = 0, \text{ for } t \in I$$

where $p(t)$ and $q(t)$ are continuous on an open interval I. Then y_1 and y_2 are linearly independent if and only if $W(y_1, y_2)(t_0) \neq 0$ for some $t_0 \in I$.

Proof. The 'if' part follows by Theorem 4.1.6. To prove the 'only if' part, suppose $W(y_1, y_2)(t) = 0$ for any $t \in I$. Take any $t_0 \in I$, we have

$$W(y_1, y_2)(t_0) = \begin{vmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{vmatrix} = 0.$$

Then system of equations

$$\left\{\begin{array}{lcl} c_1y_1(t_0)+c_2y_2(t_0) & = & 0 \\ c_1y_1'(t_0)+c_2y_2'(t_0) & = & 0 \end{array}\right. ,$$

has non-trivial solution for c_1, c_2. Now the function $c_1y_1+c_2y_2$ is a solution to the initial value problem

$$\left\{\begin{array}{l} y''+p(t)y'+q(t)y=0, \quad t\in I, \\ y(t_0)=0,\ y'(t_0)=0. \end{array}\right.$$

This initial value problem has a solution $y(t)\equiv 0$ which is unique by Theorem 4.1.1. Thus $c_1y_1+c_2y_2$ is identically zero and therefore y_1, y_2 are linearly dependent. □

Theorem 4.1.13. *Let y_1 and y_2 be solutions to*

$$L[y]=y''+p(t)y'+q(t)y=0,\ t\in I$$

where p and q are continuous on an open interval I. Then $W(y_1,y_2)(t_0)\neq 0$ for some $t_0\in I$ if and only if every solution of the equation is of the form $y=c_1y_1+c_2y_2$ for some constants c_1, c_2.

Proof. Suppose $W(y_1,y_2)(t_0)\neq 0$ for some $t_0\in I$. Let $y=y(t)$ be a solution of of $L[y]=0$ and write $y_0=y(t_0)$, $y_0'=y'(t_0)$. Since $W(t_0)\neq 0$, there exists constants c_1, c_2 such that

$$\begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}=\begin{pmatrix} y_0 \\ y_0' \end{pmatrix}.$$

Now both y and $c_1y_1+c_2y_2$ are solution to the initial problem

$$\left\{\begin{array}{l} y''+p(t)y'+q(t)y=, \quad t\in I, \\ y(t_0)=y_0, \quad y'(t_0)=y_0'. \end{array}\right.$$

Therefore $y=c_1y_1+c_2y_2$ by the uniqueness part of Theorem 4.1.1.

Suppose the general solution of $L[y]=0$ is $y=c_1y_1+c_2y_2$. Take any $t_0\in I$. Let u_1 and u_2 be solutions of $L[y]=0$ with initial values

$$\left\{\begin{array}{lcl} u_1(t_0) & = & 1 \\ u_1'(t_0) & = & 0 \end{array}\right. \text{ and } \left\{\begin{array}{lcl} u_2(t_0) & = & 0 \\ u_2'(t_0) & = & 1 \end{array}\right. .$$

The existence of u_1 and u_2 is guaranteed by Theorem 4.1.1. Thus there exists constants $a_{11}, a_{12}, a_{21}, a_{22}$ such that

$$\left\{\begin{array}{lcl} u_1 & = & a_{11}y_1+a_{21}y_2 \\ u_2 & = & a_{12}y_1+a_{22}y_2 \end{array}\right. .$$

In particular, we have

$$\left\{\begin{array}{lclcl} 1 & = & u_1(t_0) & = & a_{11}y_1(t_0)+a_{21}y_2(t_0) \\ 0 & = & u_2(t_0) & = & a_{12}y_1(t_0)+a_{22}y_2(t_0) \end{array}\right.$$

and

$$\left\{\begin{array}{lclcl} 0 & = & u_1'(t_0) & = & a_{11}y_1'(t_0)+a_{21}y_2'(t_0) \\ 1 & = & u_2'(t_0) & = & a_{12}y_1'(t_0)+a_{22}y_2'(t_0) \end{array}\right. .$$

In other words,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Therefore the matrix

$$\begin{pmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{pmatrix}$$

is non-singular and its determinant $W(y_1,y_2)(t_0)$ is non-zero. □

Combining Abel's theorem (Theorem 4.1.11), Theorem 4.1.12, Theorem 4.1.13 and the definition of basis for vector space, we obtain the following theorem.

Theorem 4.1.14. *The solution space of the second order homogeneous equation*

$$L[y] = y'' + p(t)y' + q(t)y = 0, \text{ for } t \in I,$$

where $p(t)$ and $q(t)$ are continuous on open interval I, is of dimension two. Let y_1 and y_2 be two solutions to $L[y] = 0$. Then the following statements are equivalent.

1. *$W(y_1, y_2)(t_0) \neq 0$ for some $t_0 \in I$.*
2. *$W(y_1, y_2)(t) \neq 0$ for all $t \in I$.*
3. *The functions y_1 and y_2 form a fundamental set of solutions, i.e., y_1 and y_2 are linearly independent.*
4. *The functions y_1, y_2 span the solution space of $L[y] = 0$. In other words, the general solution to the equation is $y = c_1y_1 + c_2y_2$.*
5. *The functions y_1 and y_2 constitute a basis for the solution space of $L[y] = 0$. In other words, every solution to $L[y] = 0$ can be expressed uniquely in the form $y = c_1y_1 + c_2y_2$, where c_1, c_2 are constants.*

Proof. The only thing we need to prove is that there exists solutions with $W(y_1, y_2)(t_0) \neq 0$ for some $t_0 \in I$. Take any $t_0 \in I$. By Theorem 4.1.1, there exists solutions y_1 and y_2 to the homogeneous equation $L[y] = 0$ with initial conditions

$$\begin{cases} y_1(t_0) &= 1 \\ y_1'(t_0) &= 0 \end{cases} \text{ and } \begin{cases} y_2(t_0) &= 0 \\ y_2'(t_0) &= 1 \end{cases}.$$

Then $W(y_1, y_2)(t_0) = \det(I) = 1 \neq 0$ and we are done. □

Exercise 4.1

1. Determine whether the following sets of functions are linearly dependent or independent.
   (a) $f(x) = 1 - 3x$; $g(x) = 2 - 6x$
   (b) $f(x) = 1 - x$; $g(x) = 1 + x$
   (c) $f(x) = e^x$; $g(x) = xe^x$
   (d) $f(x) = e^x$; $g(x) = 3e^{x+1}$
   (e) $f(x) = \cos x + \sin x$; $g(x) = \cos x - \sin x$
   (f) $f(x) = \cos^2 x$; $g(x) = 1 + \cos 2x$
   (g) $f(x) = \sin x \cos x$; $g(x) = \sin 2x$
   (h) $f(x) = 1 - x$; $g(x) = x - x^2$; $h(x) = x^2 - 1$
2. Find the Wronskian of the following pair of functions.
   (a) e^{2t}, e^{-3t}
   (b) $\cos t$, $\sin t$
   (c) e^{2t}, te^{2t}
   (d) t, te^t
   (e) $e^t \cos t$, $e^t \sin t$
   (f) $1 - \cos 2t$, $\sin^2 t$

3. If y_1 and y_2 form a fundamental set of solutions of $ty''+2y'+te^ty=0$ and if $W(y_1,y_2)(1)=3$, find the value of $W(y_1,y_2)(5)$.
4. If y_1 and y_2 form a fundamental set of solutions of $t^2y''-2y'+(3+t)y=0$ and if $W(y_1,y_2)(2)=3$, find the value of $W(y_1,y_2)(6)$.
5. Let $y_1(t)=t^3$ and $y_2(t)=|t|^3$. Show that $W[y_1,y_2](t)\equiv 0$. Explain why y_1 and y_2 cannot be two solutions to a homogeneous second order linear equation $y''+p(t)y'+q(t)y=0$.
6. Suppose f, g and h are differentiable functions. Show that $W(fg,fh)=f^2W(g,h)$.

4.2 Reduction of order

We have seen in the last section that to find the general solution of the homogeneous equation

$$L[y]=y''+p(t)y'+q(t)y=0,\ t\in I,$$

it suffices to find two linearly independent solutions. Suppose we know one non-zero solution $y_1(t)$ to the equation $L[y]=0$, how do we find a second solution $y_2(t)$ so that y_1 and y_2 are linearly independent? We may use the so called reduction of order. We let

$$y(t)=y_1(t)v(t).$$

where $v(t)$ is a function to be determined. Then we have

$$\begin{cases} y'=y_1v'+y_1'v, \\ y''=y_1v''+2y_1'v'+y_1''v. \end{cases}$$

Substituting them into the equation $L[y]=0$, we obtain

$$\begin{aligned} (y_1v''+2y_1'v'+y_1''v)+p(y_1v'+y_1'v)+qy_1v &= 0 \\ y_1v''+(2y_1'+py_1)v'+(y_1''+py_1'+qy_1)v &= 0 \end{aligned}$$

Since y_1 is a solution to $L[y]=0$, the coefficient of v is zero, and so the equation becomes

$$y_1v''+(2y_1'+py_1)v'=0,$$

which is a first order linear equation of v'. We can get a second solution to $L[y]=0$ by finding a non-constant solution to this first order linear order. Then we can write down the general solution to the equation $L[y]=0$.

Example 4.2.1. *Given that $y_1(t)=e^{-2t}$ is a solution to*

$$y''+4y'+4y=0,$$

find the general solution of the equation.

Solution: We set $y=e^{-2t}v$, then

$$\begin{cases} y'=e^{-2t}v'-2e^{-2t}v, \\ y''=e^{-2t}v''-4e^{-2t}v'+4e^{-2t}v. \end{cases}.$$

Thus the equation becomes

$$\begin{aligned} e^{-2t}v''-4e^{-2t}v'+4e^{-2t}v+4(e^{-2t}v'-2e^{-2t}v)+4e^{-2t}v &= 0 \\ e^{-2t}v'' &= 0 \\ v'' &= 0 \\ v' &= c_1 \\ v &= c_1t+c_2 \end{aligned}$$

Therefore the general solution is $y=e^{-2t}(c_1t+c_2)=c_1te^{-2t}+c_2e^{-2t}$. □

Example 4.2.2. *Given that* $y_1(t) = t^{-1}$ *is a solution of*

$$2t^2y'' + 3ty' - y = 0,\ t > 0,$$

find the general solution of the equation.

Solution: We set $y = t^{-1}v$, then

$$\begin{cases} y' = t^{-1}v' - t^{-2}v, \\ y'' = t^{-1}v'' - 2t^{-2}v' + 2t^{-3}v. \end{cases}$$

Thus the equation becomes

$$\begin{aligned} 2t^2(t^{-1}v'' - 2t^{-2}v' + 2t^{-3}v) + 3t(t^{-1}v' - t^{-2}v) - t^{-1}v &= 0 \\ 2tv'' - v' &= 0 \\ t^{-\frac{1}{2}}v'' - \frac{1}{2}t^{-\frac{3}{2}}v' &= 0 \\ \frac{d}{dt}(t^{-\frac{1}{2}}v') &= 0 \\ t^{-\frac{1}{2}}v' &= c \\ v' &= ct^{\frac{1}{2}} \\ v &= c_1t^{\frac{3}{2}} + c_2 \end{aligned}$$

Therefore the general solution is $y = (c_1t^{\frac{3}{2}} + c_2)t^{-1} = c_1t^{\frac{1}{2}} + c_2t^{-1}$. □

Exercise 4.2

1. Using the given solution $y_1(t)$ and the method of reduction of order to find the general solution to the following second order linear equations.

   (a) $t^2y'' - 2y = 0;\ y_1(t) = t^2$

   (b) $t^2y'' + 4ty' + 2y = 0;\ y_1(t) = t^{-1}$

   (c) $t^2y'' - ty' + y = 0;\ y_1(t) = t$

   (d) $y'' - 2y' + y = 0;\ y_1(t) = e^t$

   (e) $ty'' - y' + 4t^3y = 0;\ y_1(t) = \cos(t^2)$

   (f) $t^2y'' + 3ty' + y = 0;\ y_1(t) = t^{-1}$

   (g) $t^2y'' - 3ty' + 4y = 0;\ y_1(t) = t^2$

   (h) $t^2y'' + ty' + y = 0;\ y_1(t) = \cos(\ln x)$

4.3 Homogeneous equations with constant coefficients

We consider homogeneous equation with constant coefficients

$$L[y] = a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + cy = 0,$$

where a, b, c are constants. The equation

$$ar^2 + br + c = 0$$

is called the **characteristic equation** of the differential equation. Let r_1, r_2 be the two (complex) roots (can be equal) of the characteristic equation. The general solution of the equation can be written in terms of r_1, r_2 is given according in the following table.

Discriminant	Nature of roots	General solution
$b^2 - 4ac > 0$	r_1, r_2 are distinct and real	$y = c_1 e^{r_1 t} + c_2 e^{r_2 t}$
$b^2 - 4ac = 0$	$r_1 = r_2$ are equal	$y = c_1 e^{r_1 t} + c_2 t e^{r_1 t}$
$b^2 - 4ac < 0$	$r_1, r_2 = \lambda \pm i\mu$ $(\mu > 0)$	$y = e^{\lambda t}(c_1 \cos(\mu t) + c_2 \sin(\mu t))$

Example 4.3.1. *Solve*

$$y'' - y' - 6y = 0.$$

Solution: Solving the characteristic equation

$$\begin{aligned} r^2 - r - 6 &= 0 \\ r &= 3, -2. \end{aligned}$$

Thus the general solution is

$$y = c_1 e^{3t} + c_2 e^{-2t}.$$

□

Example 4.3.2. *Solve the initial value problem*

$$\begin{cases} y'' - 4y' + 4y = 0, \\ y(0) = 3, \ y'(0) = 1. \end{cases}$$

Solution: The characteristic equation

$$r^2 - 4r + 4 = 0$$

has a double root $r_1 = r_2 = 2$. Thus the general solution is

$$y = c_1 e^{2t} + c_2 t e^{2t}.$$

Now

$$\begin{aligned} y' &= 2c_1 e^{2t} + c_2 e^{2t} + 2c_2 t e^{2t} \\ &= (2c_1 + c_2) e^{2t} + 2c_2 t e^{2t} \end{aligned}$$

Thus

$$\begin{cases} y(0) &= c_1 = 3 \\ y'(0) &= 2c_1 + c_2 = 1 \end{cases} \Rightarrow \begin{cases} c_1 &= 3 \\ c_2 &= -5 \end{cases}$$

Therefore

$$y = 3e^{2t} - 5te^{2t}.$$

□

Example 4.3.3. *Solve the initial value problem*

$$\begin{cases} y'' - 6y' + 25y = 0, \quad t \in I \\ y(0) = 3, \ y'(0) = 1 \end{cases}.$$

Solution: The roots of the characteristic equation is

$$r_1, r_2 = 3 \pm 4i.$$

Thus the general solution is

$$y = e^{3t}(c_1 \cos 4t + c_2 \sin 4t).$$

Now

$$\begin{aligned} y' &= 3e^{3t}(c_1\cos 4t + c_2\sin 4t) + e^{3t}(-4c_1\sin 4t + 4c_2\cos 4t) \\ &= e^{3t}((3c_1+4c_2)\cos 4t + (3c_2-4c_1)\sin 4t) \end{aligned}$$

Thus

$$\left\{ \begin{array}{rcl} y(0) &=& c_1 = 3 \\ y'(0) &=& 3c_1 + 4c_2 = 1 \end{array} \right. \Rightarrow \left\{ \begin{array}{rcl} c_1 &=& 3 \\ c_2 &=& -2 \end{array} \right.$$

Therefore

$$y = e^{3t}(3\cos 4t - 2\sin 4t).$$

□

Example 4.3.4. *Solve*

$$y'' + 9y = 0.$$

Solution: The roots of the characteristic equation are $\pm 3i$. Therefore the general solution is

$$y = c_1\cos 3t + c_2\sin 3t.$$

□

One final point before we end this section. In the second case i.e. $r_1 = r_2$, the solution $y = te^{r_1 t}$ can be obtained by the method of order reduction explained in Section 4.2. Suppose the characteristic equation $ar^2 + br + c = 0$ has real and repeated roots $r_1 = r_2$. Then $y_1 = e^{r_1 t}$ is a solution to the differential equation. To find the second solution, let $y(t) = v(t)e^{r_1 t}$. Then

$$\left\{ \begin{array}{l} y' = (v' + r_1 v)e^{r_1 t} \\ y'' = (v'' + 2r_1 v' + r_1^2 v)e^{r_1 t}. \end{array} \right.$$

The equation reads

$$\begin{aligned} ay'' + by' + cy &= 0 \\ a(v'' + 2r_1 v' + r_1^2 v)e^{r_1 t} + b(v' + r_1 v)e^{r_1 t} + cve^{r_1 t} &= 0 \\ av'' + (2ar_1 + b)v' + (ar_1^2 + br_1 + c)v &= 0 \\ v'' &= 0. \end{aligned}$$

Note that r_1 is a double root of the characteristic equation, so we have $ar_1^2 + br_1 + c = 0$ and $2ar_1 + b = 0$. Hence $v(t) = c_1 + c_2 t$ for some constants c_1, c_2 and we obtain the general solution

$$y = (c_1 + c_2 t)e^{r_1 t}.$$

Exercise 4.3

1. Find the general solution of the following second order linear equations.

   (a) $y'' + y' - 6y = 0$

   (b) $y'' + 9y = 0$

   (c) $y'' - 3y' + 2y = 0$

   (d) $y'' - 8y' + 16y = 0$

   (e) $y'' + 4y' + 13y = 0$

   (f) $y'' - 2y' + 5y = 0$

2. Solve the following initial value problems.

(a) $y'' + 3y' + 2y = 0$; $y(0) = 1$, $y'(0) = 1$

(b) $y'' + 3y' = 0$; $y(0) = -2$, $y'(0) = 3$

(c) $y'' + 5y' + 6y = 0$; $y(0) = 2$, $y'(0) = 3$

(d) $4y'' + 4y' + 5y = 0$; $y(0) = 4$, $y'(0) = 1$

3. Use the substitution $u = \ln t$ to find the general solutions of the following Euler equations.

(a) $t^2y'' + 2ty' - 12y = 0$

(b) $t^2y'' - 3ty' + 4y = 0$

4.4 Method of undetermined coefficients

To solve the nonhomogeneous equation

$$L[y] = ay'' + by' + cy = g(t), \ t \in I,$$

where a, b, c are constants and $g(t)$ is a continuous function, we may first find a (particular) solution $y_p = y_p(t)$. Then the general solution is

$$y = y_c + y_p,$$

where

$$y_c = c_1y_1 + c_2y_2,$$

where y_c, which is called a complementary function, is any solution to the associated homogeneous equation $L[y] = 0$. This is because if $y = y(t)$ is a solution to $L[y] = g(t)$, then $y - y_p$ must be a solution to the associated homogeneous equation $L[y] = 0$.

When $g(t) = a_1g_1(t) + a_2g_2(t) + \cdots + a_kg_k(t)$ where $a_1, a_2, \cdots, a_k$ are real numbers and each $g_i(t)$ is of the form $e^{\alpha t}$, $\cos\omega t$, $\sin\omega t$, $e^{\alpha t}\cos\omega t$, $e^{\alpha t}\sin\omega t$, a polynomial in t or a product of a polynomial and one of the above functions, then a particular solution $y_p(t)$ is of the form which is listed in the following table.

The particular solution of $ay'' + by' + cy = g(t)$

$g(t)$	$y_p(t)$
$P_n(t) = a_nt^n + \cdots + a_1t + a_0$	$t^s(A_nt^n + \cdots + A_1t + A_0)$
$P_n(t)e^{\alpha t}$	$t^s(A_nt^n + \cdots + A_1t + A_0)e^{\alpha t}$
$P_n(t)\cos\omega t,\ P_n(t)\sin\omega t$	$t^s\left(\begin{array}{c}(A_nt^n + \cdots + A_1t + A_0)\cos\omega t\\ +(B_nt^n + \cdots + B_1t + B_0)\sin\omega t\end{array}\right)$
$P_n(t)e^{\alpha t}\cos\omega t,\ P_n(t)e^{\alpha t}\sin\omega t$	$t^se^{\alpha t}\left(\begin{array}{c}(A_nt^n + \cdots + A_1t + A_0)\cos\omega t\\ +(B_nt^n + \cdots + B_1t + B_0)\sin\omega t\end{array}\right)$

Notes: Here $s = 0, 1, 2$ is the smallest nonnegative integer that will ensure that no term in $y_p(t)$ is a solution to the associated homogeneous equation $L[y] = 0$.

The values of the constants $A_0, A_1, \cdots, A_n, B_0, B_1, \cdots, B_n$ can be obtained by substituting $y_p(t)$ to the equation $L[y] = g(t)$.

Example 4.4.1. *Find a particular solution to $y'' - y' + 2y = 4t^2$.*

Solution: A particular solution is of the form $y_p = A_2t^2 + A_1t + A_0$. To find A_0, A_1, A_2, we have

$$\begin{cases} y_p' = 2A_2t + A_1, \\ y_p'' = 2A_2. \end{cases}$$

Putting them into the equation, we have

$$\begin{aligned} y_p'' - y_p' + 2y_p &= 4t^2 \\ 2A_2 - (2A_2t + A_1) + 2(A_2t^2 + A_1t + A_0) &= 4t^2 \\ 2A_2t^2 + (-2A_2 + 2A_1)t + (2A_2 - A_1 + 2A_0) &= 4t^2 \end{aligned}$$

By comparing the coefficients, we obtain

$$\left\{ \begin{array}{rcl} 2A_2 &=& 4 \\ -2A_2 + 2A_1 &=& 0 \\ 2A_2 - A_1 + 2A_0 &=& 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{rcl} A_2 &=& 2 \\ A_1 &=& 2 \\ A_0 &=& -1 \end{array} \right. .$$

Hence a particular solution is $y_p = 2t^2 + 2t - 1$. □

Example 4.4.2. *Solve* $y'' - 3y' - 4y = 18e^{2t}$.

Solution: The roots of the characteristic equation $r^2 - 3r - 4 = 0$ is $r = 4, -1$. So the complementary function is

$$y_c = c_1e^{4t} + c_2e^{-t}.$$

Since 2 is not a root of $r^2 - 3r - 4 = 0$, we let $y_p = Ae^{2t}$, where A is a constant to be determined. Now

$$\left\{ \begin{array}{l} y_p' = 2Ae^{2t} \\ y_p'' = 4Ae^{2t} \end{array} \right.$$

By comparing coefficients of

$$\begin{aligned} y_p'' - 3y_p' - 4y_p &= 18e^{2t} \\ (4A - 3(2A) - 4A)e^{2t} &= 18e^{2t} \\ -6Ae^{2t} &= 18e^{2t} \end{aligned}$$

we get $A = -3$ and a particular solution is $y_p = -3e^{-3t}$. Therefore the general solution is

$$y = y_c + y_p = c_1e^{4t} + c_2e^{-t} - 3e^{-3t}.$$

□

Example 4.4.3. *Find a particular solution of* $y'' - 3y' - 4y = 34\sin t$.

Solution: Since $\pm i$ are not roots of $r^2 - 3r - 4 = 0$, we let

$$y_p = A\cos t + B\sin t.$$

Then

$$\left\{ \begin{array}{l} y_p' = B\cos t - A\sin t \\ y_p'' = -A\cos t - B\sin t \end{array} \right.$$

By comparing the coefficients of

$$\begin{aligned} y_p'' - 3y_p' - 4y_p &= 34\sin t \\ (-A\cos t - B\sin t) - 3(B\cos t - A\sin t) - 4(A\cos t + B\sin t) &= 34\sin t \\ (-A - 3B - 4A)\cos t + (-B + 3A - 4B)\sin t &= 34\sin t, \end{aligned}$$

we have

$$\left\{ \begin{array}{rcl} -A - 3B - 4A &=& 0 \\ -B + 3A - 4B &=& 34 \end{array} \right. \Rightarrow \left\{ \begin{array}{rcl} A &=& 3 \\ B &=& -5 \end{array} \right. .$$

Hence a particular solution is $y_p = 3\cos t - 5\sin t$. □

Example 4.4.4. *Find a particular solution of* $y'' - 3y' - 4y = 52e^t \sin 2t$.

Solution: Since $1 \pm 2i$ are not roots of $r^2 - 3r - 4 = 0$, we let

$$y_p = e^t(A\cos 2t + B\sin 2t).$$

Then

$$\begin{cases} y_p' = e^t((A+2B)\cos 2t + (B-2A)\sin 2t), \\ y_p'' = e^t((-3A+4B)\cos 2t + (-4A-3B)\sin 2t). \end{cases}$$

By comparing coefficients

$$\begin{aligned} y_p'' - 3y_p' - 4y_p &= 52e^t \sin 2t \\ e^t \left[\begin{array}{c} ((-3A+4B) - 3(A+2B) - 4A)\cos 2t \\ +((-4A-3B) - 3(B-2A) - 4B)\sin 2t \end{array} \right] &= 52e^t \sin 2t \\ (-10A - 2B)\cos 2t + (2A - 10B)\sin 2t &= 52\sin 2t \end{aligned}$$

we have $(A, B) = (1, -5)$ and a particular solution is $y_p = e^t(\cos 2t - 5\sin 2t)$. □

Example 4.4.5. *Find a particular solution of* $y'' - 3y' - 4y = 10e^{-t}$.

Solution: Since -1 is a (simple) root of the characteristic equation $r^2 - 3r - 4 = 0$, we let

$$y_p = Ate^{-t}.$$

Then

$$\begin{cases} y_p' = (-At + A)e^{-t} \\ y_p'' = (At - 2A)e^{-t}. \end{cases}$$

Now we want

$$\begin{aligned} y_p'' - 3y_p' + 4y_p &= 10e^{-t} \\ ((At - 2A) - 3(-At + A) - 4At)e^{-t} &= 10e^{-t} \\ -5Ae^{-t} &= 10e^{-t}. \end{aligned}$$

Hence we take $A = -2$ and a particular solution is $y_p = -2te^{-t}$. □

Example 4.4.6. *Find a particular solution of* $y'' - 3y' - 4y = 10e^{-t} + 34\sin t + 52e^t \sin 2t$.

Solution: From the above three examples, a particular solution is

$$y_p = -2te^{-t} + 3\cos t - 5\sin t + e^t\cos 2t - 5e^t \sin 2t.$$

□

Example 4.4.7. *Find a particular solution of* $y'' + 4y = 4\cos 2t$.

Solution: Since $\pm i$ are roots of the characteristic equation $r^2 + 4 = 0$, we let

$$y_p = At\cos 2t + Bt\sin 2t.$$

Then

$$\begin{cases} y_p' = (2Bt + A)\cos 2t + (-2At + B)\sin 2t \\ y_p'' = (-4At + 4B)\cos 2t + (-4Bt - 4A)\sin 2t. \end{cases}$$

By comparing coefficients of

$$\begin{aligned} y_p'' + 4y_p &= 4\cos 2t \\ (-4At + 4B)\cos 2t + (-4Bt - 4A)\sin 2t + 4(At\cos 2t + Bt\sin 2t) &= 4\cos 2t \\ 4B\cos 2t - 4A\sin 2t &= 4\cos 2t, \end{aligned}$$

we take $A = 0$, $B = 1$ and a particular solution is $y_p = t\cos 2t$. □

Example 4.4.8. *Solve* $y'' + 2y' + y = 6te^{-t}$.

Solution: The characteristic equation $r^2+2r+1 = 0$ has a double root -1. So the complementary function is

$$y_c = c_1e^{-t} + c_2te^{-t}.$$

Since -1 is a double root of the characteristic equation, we let $y_p = t^2(At+B)e^{-t}$, where A and B are constants to be determined. Now

$$\begin{cases} y_p' = (-At^3 + (3A - B)t^2 + 2Bt)e^{-t} \\ y_p'' = (At^3 + (-6A + B)t^2 + (6A - 4B)t + 2B)e^{-t}. \end{cases}$$

By comparing coefficients of

$$\begin{aligned} y_p'' + 2y_p' + y_p &= 6te^{-t} \\ \begin{bmatrix} (At^3 + (-6A + B)t^2 + (6A - 4B)t + 2B) \\ +2(-At^3 + (3A - B)t^2 + 2Bt) \\ +(At^3 + Bt^2) \end{bmatrix} e^{-t} &= 6te^{-t} \\ (6At + 2B)e^{-t} &= 6te^{-t}, \end{aligned}$$

we take $A = 1$, $B = 0$ and a particular solution is $y_p = t^3e^{-t}$. Therefore the general solution is

$$y = y_c + y_p = c_1e^{-t} + c_2te^{-t} + t^3e^{-t}.$$

□

Example 4.4.9. *Determine the appropriate form for a particular solution of*

$$y'' + y' - 2y = 3t - \sin 4t + 3t^2e^{2t}.$$

Solution: The characteristic equation $r^2 + r - 2 = 0$ has roots $r = 2, -1$. So the complementary function is

$$y_c = c_1e^{2t} + c_2e^{-t}.$$

A particular solution takes the form

$$y_p = (A_1t + A_0) + (B_1 \cos 4t + B_2 \sin 4t) + t(C_2t^2 + C_1t + C_0)e^{2t}.$$

□

Example 4.4.10. *Determine the appropriate form for a particular solution of*

$$y'' + 2y' + 5y = te^{3t} - t\cos t + 2te^{-t} \sin 2t.$$

Solution: The characteristic equation $r^2+2r+5 = 0$ has roots $r = -1\pm 2i$. So the complementary function is

$$y_c = e^{-t}(c_1 \cos 2t + c_2 \sin 2t).$$

A particular solution takes the form

$$y_p = (A_1t+A_0)e^{3t}+(B_1t+B_2)\cos t+(B_3t+B_4)\sin t+te^{-t}((C_1t+C_2)\cos 2t+(C_3t+C_4)\sin 2t).$$

□

Exercise 4.4

1. Use the method of undetermined coefficients to find the general solution of the following nonhomogeneous second order linear equations.

(a) $y'' - 2y' - 3y = 3e^{2t}$

(b) $y'' + 2y' + 5y = 3\cos 2t$

(c) $y'' + 9y = t^2e^{3t} + 6$

(d) $y'' + y' - 2y = 2t$

(e) $y'' + 2y' + y = 2e^{-t}$

(f) $y'' - 2y' + y = te^t + 4$

(g) $y'' + 4y = t^2 + 3e^t$

(h) $y'' + 2y' + 5y = 4e^{-t}\cos 2t$

(i) $y'' - 3y' - 4y = 6e^{2t} + 2\sin t$

(j) $y'' + 4y' + 4y = 8e^{2t} + 8e^{-2t}$

2. Write down a suitable form $y_p(t)$ of a particular solution of the following nonhomogeneous second order linear equations.

(a) $y'' + 3y' = 2t^4 + t^2e^{-3t} + \sin 3t$

(b) $y'' - 5y' + 6y = e^t\cos 2t + 3te^{2t}\sin t$

(c) $y'' + y = t(1 + \sin t)$

(d) $y'' + 2y' + 2y = e^{-t}(2 - \cos t + 5t^2\sin t)$

(e) $y'' - 4y' + 4y = t(2 + e^{2t} - 3\cos 2t)$

4.5 Variation of parameters

To solve a non-homogeneous equation $L[y] = g(t)$, we may use another method called variation of parameters. Suppose $y(t) = c_1y_1(t)+c_2y_2(t)$, where c_1, c_2 are constants, is the general solution to the associated homogeneous equation $L[y] = 0$. The idea is to let the two parameters c_1, c_2 vary and see whether we could choose suitable functions $u_1(t), u_2(t)$, which depend on $g(t)$, such that $y = u_1(t)y_1(t) + u_2(t)y_2(t)$ is solution to $L[y] = g(t)$. It turns out that this is always possible.

Theorem 4.5.1. *Let y_1 and y_2 be a fundamental set of solution of the homogeneous equation*

$$L[y] = y'' + p(t)y' + q(t)y = 0,\ t \in I,$$

where $p(t)$ and $q(t)$ are continuous functions on I and

$$W = W(y_1, y_2)(t) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1y_2' - y_1'y_2,$$

be the Wronskian. Let $g = g(t)$ be any continuous function. Suppose u_1 and u_2 are differentiable functions on I such that

$$\begin{cases} u_1' = -\dfrac{gy_2}{W} \\ u_2' = \dfrac{gy_1}{W}, \end{cases}$$

then $y_p = u_1y_1 + u_2y_2$ is a solution to

$$L[y] = g,\ t \in I.$$

Proof. Observe that u_1' and u_2' are solution to the system of equations

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} 0 \\ g \end{pmatrix}$$

or equivalently

$$\begin{cases} u_1'y_1 + u_2'y_2 &= 0 \\ u_1'y_1' + u_2'y_2' &= g \end{cases}.$$

Hence we have

$$\begin{aligned} y_p' &= u_1'y_1 + u_2'y_2 + u_1y_1' + u_2y_2' \\ &= u_1y_1' + u_2y_2' \end{aligned}$$

and

$$\begin{aligned} y_p'' &= u_1'y_1' + u_2'y_2' + u_1y_1'' + u_2y_2'' \\ &= g(t) + u_1y_1'' + u_2y_2''. \end{aligned}$$

Therefore

$$\begin{aligned} y_p'' + p(t)y_p' + q(t)y_p &= g(t) + u_1y_1'' + u_2y_2'' + p(t)(u_1y_1' + u_2y_2') + q(t)(u_1y_1 + u_2y_2) \\ &= g(t) + u_1(y_1'' + p(t)y_1' + q(t)y_1) + u_2(y_2'' + p(t)y_2' + q(t)y_2) \\ &= g(t) \end{aligned}$$

where the last equality follows from the fact that y_1 and y_2 are solution to $L[y] = 0$. □

Example 4.5.2.

$$y'' + 4y = \frac{3}{\sin t}$$

Solution: Solving the corresponding homogeneous equation, we let

$$y_1 = \cos 2t,\ \ y_2 = \sin 2t.$$

We have

$$W(y_1, y_2)(t) = \begin{vmatrix} \cos 2t & \sin 2t \\ -2\sin 2t & 2\cos 2t \end{vmatrix} = 2.$$

So

$$\begin{cases} u_1' = -\dfrac{gy_2}{W} = -\dfrac{\left(\frac{3}{\sin t}\right)\sin 2t}{2} = -\dfrac{3(2\cos t\sin t)}{2\sin t} = -3\cos t, \\ u_2' = \dfrac{gy_1}{W} = \dfrac{\left(\frac{3}{\sin t}\right)\cos 2t}{2} = \dfrac{3(1 - 2\sin^2 t)}{2\sin t} = \dfrac{3}{2\sin t} - 3\sin t. \end{cases}$$

Hence,

$$\begin{cases} u_1 = -3\sin t + c_1, \\ u_2 = \frac{3}{2}\ln|\csc t - \cot t| + 3\cos t + c_2. \end{cases}$$

and the general solution is

$$\begin{aligned} y &= u_1y_1 + u_2y_2 \\ &= (-3\sin t + c_1)\cos 2t + (\frac{3}{2}\ln|\csc t - \cot t| + 3\cos t + c_2)\sin 2t \\ &= c_1\cos 2t + c_2\sin 2t - 3\sin t\cos 2t + \frac{3}{2}\sin 2t\ln|\csc t - \cot t| + 3\cos t\sin 2t \\ &= c_1\cos 2t + c_2\sin 2t + \frac{3}{2}\sin 2t\ln|\csc t - \cot t| + 3\sin t \end{aligned}$$

where c_1, c_2 are constants. □

Example 4.5.3.

$$y'' - 3y' + 2y = \frac{e^{3t}}{e^t + 1}$$

Solution: Solving the corresponding homogeneous equation, we let

$$y_1 = e^t,\ y_2 = e^{2t}.$$

We have

$$W(y_1, y_2)(t) = \begin{vmatrix} e^t & e^{2t} \\ e^t & 2e^{2t} \end{vmatrix} = e^{3t}.$$

So

$$\begin{cases} u_1' = -\dfrac{gy_2}{W} = -\left(\dfrac{e^{3t}}{e^t+1}e^{2t}\right)/e^{3t} = -\dfrac{e^{2t}}{e^t+1}, \\ u_2' = \dfrac{gy_1}{W} = \left(\dfrac{e^{3t}}{e^t+1}e^t\right)/e^{3t} = \dfrac{e^t}{e^t+1}. \end{cases}$$

Thus

$$\begin{aligned} u_1 &= -\int \frac{e^{2t}}{e^t+1}dt \\ &= -\int \frac{e^t}{e^t+1}de^t \\ &= \int \left(\frac{1}{e^t+1} - 1\right) d(e^t+1) \\ &= \log(e^t+1) - (e^t+1) + c_1 \end{aligned}$$

and

$$\begin{aligned} u_2 &= \int \frac{e^t}{e^t+1}dt \\ &= \int \frac{1}{e^t+1}d(e^t+1) \\ &= \log(e^t+1) + c_2. \end{aligned}$$

Therefore the general solution is

$$\begin{aligned} y &= u_1y_1 + u_2y_2 \\ &= (\log(e^t+1) - (e^t+1) + c_1)e^t + (\log(e^t+1) + c_2)e^{2t} \\ &= c_1e^t + c_2e^{2t} + (e^t + e^{2t})\log(e^t+1) \end{aligned}$$

□

Exercise 4.5

1. Use the method of variation of parameters to find the general solution of the following nonhomogeneous second order linear equations.
   (a) $y'' - 5y' + 6y = 2e^t$
   (b) $y'' - y' - 2y = 2e^{-t}$
   (c) $y'' + 2y' + y = 4e^{-t}$
   (d) $y'' + y = \tan t;\ 0 < t < \frac{\pi}{2}$
   (e) $y'' + 9y = 9\sec^2 3t;\ 0 < t < \frac{\pi}{6}$
   (f) $y'' - 2y' + y = \dfrac{e^t}{1+t^2}$
   (g) $y'' - 3y' + 2y = \dfrac{1}{1+e^{-t}}$

2. Use the method of variation of parameters to find the general solution of the following nonhomogeneous second order linear equations. Two linearly independent solutions $y_1(t)$ and $y_2(t)$ to the corresponding homogeneous equations are given.

(a) $t^2y'' - 2y = 3t^2 - 1$, $t > 0$; $y_1(t) = t^{-1}$, $y_2(t) = t^2$

(b) $t^2y'' - t(t+2)y' + (t+2)y = 2t^3$, $t > 0$; $y_1(t) = t$, $y_2(t) = te^t$

(c) $(1-t)y'' + ty' - y = 2(t-1)^2e^{-t}$, $0 < t < 1$; $y_1(t) = t$, $y_2(t) = e^t$

(d) $t^2y'' - 3ty' + 4y = t^2 \ln t$, $t > 0$; $y_1(t) = t^2$, $y_2(t) = t^2 \ln t$

4.6 Mechanical and electrical vibrations

One of the reasons why second order linear equations with constant coefficients are worth studying is that they serve as mathematical models of simple vibrations.

Mechanical vibrations

Consider a mass m hanging on the end of a vertical spring of original length l. Let $u(t)$, measured positive downward, denote the displacement of the mass from its equilibrium position at time t. Then $u(t)$ is related to the forces acting on the mass through Newton's law of motion

$$mu''(t) + ku(t) = f(t), \tag{2.6.1}$$

where k is the spring constant and $f(t)$ is the net force (excluding gravity and force from the spring) acting on the mass.

Undamped free vibrations

If there is no external force, then $f(t) = 0$ and equation (2.6.1) reduces to

$$mu''(t) + ku(t) = 0.$$

The general solution is

$$u = C_1 \cos \omega_0 t + C_2 \sin \omega_0 t,$$

where

$$\omega_0 = \sqrt{\frac{k}{m}}$$

is the **natural frequency** of the system. The **period** of the vibration is given by

$$T = \frac{2\pi}{\omega_0} = 2\pi\sqrt{\frac{m}{k}}.$$

We can also write the solution in the form

$$u(t) = A\cos(\omega_0 t - \alpha).$$

Then A is the **amplitude** of the vibration. Moreover, u satisfies the initial conditions

$$u(0) = u_0 = A\cos\alpha \text{ and } u'(0) = u_0' = A\omega_0 \sin\alpha.$$

Thus we have

$$A = u_0^2 + \frac{u_0'^2}{\omega_0^2} \text{ and } \alpha = \tan^{-1}\frac{u_0'}{u_0\omega_0}.$$

Damped free vibrations

If we include the effect of damping, the differential equation governing the motion of mass is

$$mu'' + \gamma u' + ku = 0,$$

where γ is the damping coefficient. The roots of the corresponding characteristic equation are

$$r_1, r_2 = \frac{-\gamma \pm \sqrt{\gamma^2 - 4km}}{2m}.$$

The solution of the equation depends on the sign of $\gamma^2 - 4km$ and are listed in the following table.

Solution of $mu'' + \gamma u' + ku = 0$

Case	Solution	Damping
$\dfrac{\gamma^2}{4km} < 1$	$e^{-\frac{\gamma}{2m}t}(C_1 \cos \mu t + C_2 \sin \mu t)$	Small damping
$\dfrac{\gamma^2}{4km} = 1$	$(C_1 t + C_2)e^{-\frac{\gamma}{2m}t}$	Critical damping
$\dfrac{\gamma^2}{4km} > 1$	$C_1 e^{r_1 t} + C_2 e^{r_2 t}$	Overdamped

Here

$$\mu = \sqrt{\frac{k}{m} - \frac{\gamma^2}{4m^2}} = \omega_0 \sqrt{1 - \frac{\gamma^2}{4km}}$$

is called the **quasi frequency**. As $\gamma^2/4km$ increases from 0 to 1, the quasi frequency μ decreases from $\omega_0 = \sqrt{k/m}$ to 0 and the quasi period increases from $2\pi\sqrt{m/k}$ to infinity.

Electric circuits

Second order linear differential equation with constant coefficients can also be used to study electric circuits. By Kirchhoff's law of electric circuit, the total charge Q on the capacitor in a simple series LCR circuit satisfies the differential equation

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = E(t),$$

where L is the inductance, R is the resistance, C is the capacitance and $E(t)$ is the impressed voltage. Since the flow of current in the circuit is $I = dQ/dt$, differentiating the equation with respect to t gets

$$LI'' + RI' + C^{-1}I = E'(t).$$

Therefore the results for mechanical vibrations in the preceding paragraphs can be used to study LCR circuit.

Forces vibrations with damping

Suppose that an external force $F_0 \cos \omega t$ is applied to a damped ($\gamma > 0$)) spring-mass system. Then the equation of motion is

$$mu'' + \gamma u' + ku = F_0 \cos \omega t.$$

The general solution of the equation must be of the form

$$u = c_1 u_1(t) + c_2 u_2(t) + A\cos(\omega t - \alpha) = u_c(t) + U(t).$$

Since m, γ, k are all positive, the real part of the roots of the characteristic equation are always negative. Thus $u_c \to 0$ as $t \to \infty$ and it is called the **transient solution**. The remaining term $U(t)$ is called the **steady-state solution** or the **forced response**. Straightforward, but somewhat lengthy computations shows that

$$A = \frac{F_0}{\Delta}, \quad \cos\alpha = \frac{m(\omega_0^2 - \omega^2)}{\Delta}, \quad \sin\alpha = \frac{\gamma\omega}{\Delta},$$

where

$$\Delta = \sqrt{m^2(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \quad \text{and} \quad \omega_0 = \sqrt{k/m}.$$

If $\gamma^2 < 2mk$, **resonance** occurs, i.e. the maximum amplitude

$$A_{\max} = \frac{F_0}{\gamma\omega_0\sqrt{1 - \gamma^2/4mk}}$$

is obtained, when

$$\omega = \omega_{\max} = \omega_0^2 \left(1 - \frac{\gamma^2}{2mk}\right).$$

We list in the following table how the amplitude A and phase angle α of the steady-state oscillation depends on the frequency ω of the external force.

Amplitude and phase of forced vibration

Frequency	Amplitude	Phase angle
$\omega \to 0$	$A \to \frac{F_0}{k}$	$\alpha \to 0$
$\omega = \omega_0^2 \left(1 - \frac{\gamma^2}{2mk}\right)$	$\frac{F_0}{\gamma\omega_0\sqrt{1-\gamma^2/4mk}}$	$\frac{\pi}{2}$
$\omega \to \infty$	$A \to 0$	$\alpha \to \pi$

Forced vibrations without damping
The equation of motion of an undamped forced oscillator is

$$mu'' + ku = F_0 \cos \omega t.$$

The general solution of the equation is

$$\begin{cases} u = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t + \dfrac{F_0 \cos \omega t}{m(\omega_0^2 - \omega^2)}, & \omega \neq \omega_0 \\ u = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t + \dfrac{F_0 t \sin \omega_0 t}{2m\omega_0}, & \omega = \omega_0 \end{cases}$$

Suppose $\omega \neq \omega_0$. If we assume that the mass is initially at rest so that the initial condition are $u(0) = u'(0) = 0$, then the solution is

$$\begin{aligned} u &= \frac{F_0}{m(\omega_0^2 - \omega^2)}(\cos \omega t - \cos \omega_0 t) \\ &= \frac{2F_0}{m(\omega_0^2 - \omega^2)} \sin \frac{(\omega_0 - \omega)t}{2} \sin \frac{(\omega_0 + \omega)t}{2}. \end{aligned}$$

If $|\omega_0 - \omega|$ is small, then $\omega_0 + \omega$ is much greater than $|\omega_0 - \omega|$. The motion is a rapid oscillation with frequency $(\omega_0 + \omega)/2$ but with a slowly varying sinusoidal amplitude

$$\frac{2F_0}{m|\omega_0^2 - \omega^2|} \left| \sin \frac{(\omega_0 - \omega)t}{2} \right|.$$

This type of motion is called a **beat** and $|\omega_0 - \omega|/2$ is the **beat frequency**.

4.7 Higher order linear equations

The theoretical structure and methods of solution developed in the preceding sections for second order linear equations extend directly to linear equations of third and higher order. Consider the nth order linear differential equation

$$L[y] = y^{(n)} + p_{n-1}(t)y^{(n-1)} + \cdots + p_1(t)y' + p_0(t)y = g(t), \ t \in I,$$

where $p_0(t), p_1(t), \cdots, p_{n-1}(t)$ are continuous functions on I. We may generalize the Wronskian as follow.

Definition 4.7.1 (Wronskian)**.** *Let* $y_1, y_2, \cdots, y_n$ *be differentiable functions. Then we define the* **Wronskian** *to be the function*

$$W = W(t) = \begin{vmatrix} y_1(t) & y_2(t) & \cdots & y_n(t) \\ y_1'(t) & y_2'(t) & \cdots & y_n'(t) \\ y_1''(t) & y_2''(t) & \cdots & y_n''(t) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & \cdots & y_n^{(n-1)}(t) \end{vmatrix}.$$

We have the analogue of Theorem 4.1.11 and Theorem 4.1.14 for higher order equation.

Theorem 4.7.2 (Abel's Theorem)**.** *Suppose* $y_1, y_2, \cdots, y_n$ *are solutions of the homogeneous equation*

$$L[y] = \frac{d^n y}{dt^n} + p_{n-1}(t)\frac{d^{n-1}y}{dt^{n-1}} \cdots + p_1(t)\frac{dy}{dt} + p_0(t)y = 0, \ \text{on } I.$$

Then the Wronskian $W(y_1, y_2, \cdots, y_n)$ *satisfies*

$$W(y_1, y_2, \cdots, y_n)(t) = c\exp\left(-\int p_{n-1}(t)dt\right)$$

for some constant c.

Theorem 4.7.3. *The solution space of the homogeneous equation*

$$L[y] = y^{(n)} + p_{n-1}(t)y^{(n-1)} + \cdots + p_1(t)y' + p_0(t)y = 0, \ t \in I,$$

is of dimension n. *Let* $y_1, y_2, \cdots, y_n$ *be solutions of* $L[y] = 0$, *then the following statements are equivalent.*

1. $W(t_0) \neq 0$ *for some* $t_0 \in I$.
2. $W(t) \neq 0$ *for all* $t \in I$.
3. *The functions* $y_1, y_2, \cdots, y_n$ *form a fundamental set of solutions, i.e.,* $y_1, y_2, \cdots, y_n$ *are linearly independent.*
4. *Every solution of the equation is of the form* $c_1y_1 + c_2y_2 + \cdots + c_ny_n$ *for some constants* $c_1, c_2, \cdots, c_n$, *i.e.,* $y_1, y_2, \cdots, y_n$ *span the solution space of* $L[y] = 0$.
5. *The functions* $y_1, y_2, \cdots, y_n$ *constitute a basis for the solution space of* $L[y] = 0$.

Now we assume that the coefficients are constants and consider

$$L[y] = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0, \ t \in I,$$

where $a_0, a_1, \cdots, a_{n-1}$ are constants. The equation

$$r^n + a_{n-1}r^{n-1} + \cdots + a_1r + a_0 = 0,$$

is called the characteristic equation of the differential equation.

If λ is a real root of the characteristic equation with multiplicity m, then

$$e^{\lambda t}, te^{\lambda t}, \cdots, t^{m-1}e^{\lambda t}$$

are solutions to the equation.

If $\pm\mu i$ are purely imaginary roots of the characteristic equation with multiplicity m, then

$$\cos\mu t, t\cos\mu t, \cdots, t^{m-1}\cos\mu t, \text{ and } \sin\mu t, t\sin\mu t, \cdots, t^{m-1}\sin\mu t$$

are solutions to the equation.

If $\lambda \pm \mu i$ are complex roots of the characteristic equation with multiplicity m, then

$$e^{\lambda t}\cos\mu t, te^{\lambda t}\cos\mu t, \cdots, t^{m-1}e^{\lambda t}\cos\mu t,$$

and

$$e^{\lambda t}\sin\mu t, te^{\lambda t}\sin\mu t, \cdots, t^{m-1}e^{\lambda t}\sin\mu t$$

are solutions to the equation.

We list the solutions for differential nature of roots in the following table.

Solutions of $L[y] = 0$

Root with multiplicity m	Solutions
Real number λ	$e^{\lambda t}, te^{\lambda t}, \cdots, t^{m-1}e^{\lambda t}$
Purely imaginary number μi	$\cos\mu t, t\cos\mu t, \cdots, t^{m-1}\cos\mu t,$ $\sin\mu t, t\sin\mu t, \cdots, t^{m-1}\sin\mu t$
Complex number $\lambda + \mu i$	$e^{\lambda t}\cos\mu t, te^{\lambda t}\cos\mu t, \cdots, t^{m-1}e^{\lambda t}\cos\mu t,$ $e^{\lambda t}\sin\mu t, te^{\lambda t}\sin\mu t, \cdots, t^{m-1}e^{\lambda t}\sin\mu t$

Note that by fundamental theorem of algebra, there are exactly n functions which are of the above forms. It can be proved that the Wronskian of these function are not identically zero. Thus these n functions constitute a fundamental set of solutions to the homogeneous equation $L[y] = 0$.

Example 4.7.4. *Find the general solution of*

$$y^{(4)} + y''' - 7y'' - y' + 6y = 0.$$

Solution: The roots of the characteristic equation

$$r^4 + r^3 - 7r^2 - r + 6 = 0$$

are

$$r = -3, -1, 1, 2.$$

Therefore the general solution is

$$y(t) = c_1e^{-3t} + c_2e^{-t} + c_3e^t + c_4e^{2t}.$$

□

Example 4.7.5. *Solve the initial value problem*

$$\begin{cases} y^{(4)} - y = 0 \\ (y(0), y'(0), y''(0), y'''(0)) = (2, -1, 0, 3). \end{cases}$$

Solution: The roots of the characteristic equation

$$r^4 - 1 = 0$$

are

$$r = \pm 1, \pm i.$$

Therefore the solution is of the form

$$y(t) = c_1 e^t + c_2 e^{-t} + c_3 \cos t + c_4 \sin t.$$

The initial condition gives

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix}$$

We find that

$$\begin{aligned} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & -1 \\ 2 & 0 & -2 & 0 \\ 0 & 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ 1 \\ -2 \end{pmatrix} \end{aligned}$$

Therefore the solution to the initial value problem is $y = e^t + \cos t - 2 \sin t$. □

Example 4.7.6. *Find the general solution of*

$$y^{(4)} + 2y'' + y = 0.$$

Solution: The characteristic equation is

$$r^4 + 2r^2 + 1 = 0$$

and its roots are

$$r = i, i, -i, -i.$$

Thus the general solution is

$$y(t) = c_1 \cos t + c_2 \sin t + c_3 t \cos t + c_4 t \sin t.$$

□

Example 4.7.7. *Find the general solution of*

$$y^{(4)} + y = 0.$$

Solution: The characteristic equation is

$$r^4 + 1 = 0$$

and its roots are

$$\begin{aligned} r &= e^{\frac{2k+1}{4}\pi i}, \quad k = 0,1,2,3 \\ &= \cos(\frac{2k+1}{4}\pi) + i\sin(\frac{2k+1}{4}\pi), \quad k = 0,1,2,3 \\ &= \pm\frac{\sqrt{2}}{2} \pm \frac{\sqrt{2}}{2} i \end{aligned}$$

Thus the general solution is

$$y(t) = c_1 e^{\frac{\sqrt{2}}{2}t}\cos\frac{\sqrt{2}}{2}t + c_2 e^{\frac{\sqrt{2}}{2}t}\sin\frac{\sqrt{2}}{2}t + c_3 e^{-\frac{\sqrt{2}}{2}t}\cos\frac{\sqrt{2}}{2}t + c_4 e^{\frac{-\sqrt{2}}{2}t}\sin\frac{\sqrt{2}}{2}t.$$

□

Method of undetermined coefficients

The main difference in using the method of undetermined coefficients for higher order equations stems from the fact that roots of the characteristic equation may have multiplicity greater than 2. Consequently, terms proposed for the nonhomogeneous part of the solution may need to be multiplied by higher powers of t to make them different from terms in the solution of the corresponding homogeneous equation.

Example 4.7.8. *Find the general solution of*

$$y''' - 3y'' + 3y' - y = 2te^t - e^t.$$

Solution: The characteristic equation

$$r^3 - 3r^2 + 3r - 1 = 0$$

has a triple root $r = 1$. So the general solution of the associated homogeneous equation is

$$y_c(t) = c_1 e^t + c_2 te^t + c_3 t^2 e^t.$$

Since $r = 1$ is a root of multiplicity 3, a particular solution is of the form

$$y_p(t) = t^3(At + B)e^t = (At^4 + Bt^3)e^t.$$

We have

$$\begin{cases} y_p' = (At^4 + (4A+B)t^3 + 3Bt^2)e^t \\ y_p'' = (At^4 + (8A+B)t^3 + (12A+6B)t^2 + 6Bt)e^t \\ y_p'' = (At^4 + (12A+B)t^3 + (36A+9B)t^2 + (24A+18B)t + 6B)e^t \end{cases}$$

Substituting $y_p(t)$ into the equation, we have

$$\begin{aligned} y_p''' - 3y_p'' + 3y_p' - y_p &= 2te^t - e^t \\ 24Ate^t + 6Be^t &= 2te^t - e^t. \end{aligned}$$

(Note that the coefficients of t^4e^t, t^3e^t, t^2e^t will automatically be zero since $r = 1$ is a triple root of the characteristic equation.) Thus

$$A = \frac{1}{12}, \quad B = -\frac{1}{6}.$$

Therefore the general solution is

$$y(t) = c_1e^t + c_2te^t + c_3t^2e^t - \frac{1}{6}t^3e^t + \frac{1}{12}t^4e^t.$$

□

Example 4.7.9. *Find a particular solution of the equation*

$$y^{(4)} + 2y'' + y = 4\cos t - \sin t.$$

Solution: The general solution of the associated homogeneous equation is

$$y_c(t) = c_1\cos t + c_2\sin t + c_3t\cos t + c_4t\sin t.$$

Since $r = \pm i$ are double roots of the equation, a particular solution is of the form

$$y_p(t) = t^2(A\cos t + B\sin t) = At^2\cos t + Bt^2\sin t.$$

We have

$$\begin{cases} y_p' = (Bt^2 + 2At)\cos t + (-At^2 + 2Bt)\sin t \\ y_p'' = (-At^2 + 4Bt + 2A)\cos t + (-Bt^2 - 4At + 2B)\sin t \\ y_p^{(3)} = (-Bt^2 - 6At + 6B)\cos t + (At^2 - 6Bt - 6A)\sin t \\ y_p^{(4)} = (At^2 - 8Bt - 12A)\cos t + (Bt^2 + 8At - 12B)\sin t \end{cases}$$

Substitute y_p into the equation, we have

$$\begin{aligned} y_p^{(4)} + 2y_p'' + y_p &= 4\cos t - \sin t \\ -8A\cos t - 8B\sin t &= 4\cos t - \sin t. \end{aligned}$$

(Note that the coefficients of $t^2\cos t$, $t^2\sin t$, $t\cos t$, $t\sin t$ will automatically be zero since $r = \pm i$ are double roots of the characteristic equation.) Thus

$$A = -\frac{1}{2},\quad B = \frac{1}{8}.$$

Therefore the general solution of the equation is

$$y(t) = c_1\cos t + c_2\sin t + c_3t\cos t + c_4t\sin t - \frac{1}{2}t^2\cos t + \frac{1}{8}t^2\sin t.$$

□

Example 4.7.10. *Find a particular solution of*

$$y''' - 9y' = t^2 + 3\sin t + e^{3t}.$$

Solution: The roots of the characteristic equation are $r = 0, \pm 3$. A particular solution is of the form

$$y_p(t) = A_1t^3 + A_2t^2 + A_3t + B_1\cos t + B_2\sin t + Cte^{3t}.$$

Substituting into the equation, we have

$$6A_1 - 9A_3 - 18A_2t - 27A_1t^2 - 10B_2\cos t + 10B_1\sin t + 18Ce^{3t} = t^2 + 3\sin t + e^{3t}.$$

Thus

$$A_1 = -\frac{1}{27},\ A_2 = 0,\ A_3 = -\frac{2}{81},\ B_1 = \frac{3}{10},\ B_2 = 0,\ C = \frac{1}{18}.$$

A particular solution is

$$y_p(t) = -\frac{1}{27}t^3 - \frac{2}{81}t + \frac{3}{10}\cos t + \frac{1}{18}te^{3t}.$$

□

Example 4.7.11. *Determine the appropriate form for a particular solution of*

$$y''' + 6y'' + 12y' + 8y = 4t - 3t^2e^{-2t} - t\sin 3t.$$

Solution: The characteristic equation $r^3 + 6r^2 + 12r + 8 = 0$ has one triple root $r = -2$. So the complementary function is

$$y_c = c_1e^{-2t} + c_2te^{-2t} + c_3t^2e^{-2t}.$$

A particular solution takes the form

$$y_p = A_1t + A_0 + t^3(B_2t^2 + B_1t + B_0)e^{-2t} + (C_1t + C_2)\cos 3t + (C_3t + C_4)\sin 3t.$$

□

Example 4.7.12. *Determine the appropriate form for a particular solution of*

$$y^{(4)} + 4y'' + 4y = 5e^t\sin 2t - 2t\cos 2t.$$

Solution: The characteristic equation $r^4 + 4r^2 + 4 = 0$ has two double roots $r = \pm 2i$. So the complementary function is

$$y_c = (c_1t + c_2)\cos 2t + (c_3t + c_4)\sin 2t).$$

A particular solution takes the form

$$y_p = e^t(A_1\cos 2t + A_2\sin 2t) + t^2((B_1t + B_2)\cos 2t + (B_3t + B_4)\sin 2t).$$

□

Method of variation of parameters

Theorem 4.7.13. *Suppose $y_1, y_2, \cdots, y_n$ are solutions of the homogeneous linear differential equation*

$$L[y] = \frac{d^ny}{dt^n} + p_{n-1}(t)\frac{d^{n-1}y}{dt^{n-1}}\cdots + p_1(t)\frac{dy}{dt} + p_0(t)y = 0, \text{ on } I.$$

Let $W(t)$ be the Wronskian and $W_k(t)$ be the determinant obtained from replacing the k-th column of $W(t)$ by the column $(0, \cdots, 0, 1)^T$. For any continuous function $g(t)$ on I, the function

$$y_p(t) = \sum_{k=1}^{n} y_k(t)\int_{t_0}^{t}\frac{g(s)W_k(s)}{W(s)}ds,$$

is a particular solution to the non-homogeneous equation

$$L[y](t) = g(t).$$

Outline of proof. Let

$$y_p(t) = v_1y_1 + v_2y_2 + \cdots + v_ny_n.$$

To choose functions $v_1, v_2, \cdots, v_n$ so that $y_p(t)$ is a solution to

$$L[y] = g,$$

we want them satisfy

$$\left\{\begin{array}{ccccccccc} v_1'y_1 & + & v_2'y_2 & + & \cdots & + & v_n'y_n & = & 0 \\ v_1'y_1' & + & v_2'y_2' & + & \cdots & + & v_n'y_n' & = & 0 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ v_1'y_1^{(n-2)} & + & v_2'y_2^{(n-2)} & + & \cdots & + & v_n'y_n^{(n-2)} & = & 0 \\ v_1'y_1^{(n-1)} & + & v_2'y_2^{(n-1)} & + & \cdots & + & v_n'y_n^{(n-1)} & = & g \end{array}\right.$$

The result is obtained by solving the above system of equations. □

Example 4.7.14. *Find a particular solution of*

$$y''' - 2y'' + y' - 2y = 5t.$$

Solution: The roots of the characteristic equation are $r = 2, \pm i$. A set of fundamental solutions is given by

$$\begin{cases} y_1 &= e^{2t} \\ y_2 &= \cos t \\ y_3 &= \sin t \end{cases}$$

The Wronskian is

$$W(t) = \begin{vmatrix} e^{2t} & \cos t & \sin t \\ 2e^{2t} & -\sin t & \cos t \\ 4e^{2t} & -\cos t & -\sin t \end{vmatrix} = 5e^{2t}$$

Using variation of parameter, let

$$y_p(t) = v_1 e^{2t} + v_2 \cos t + v_3 \sin t,$$

where v_1, v_2, v_3 satisfy

$$\begin{cases} v_1' e^{2t} & + & v_2' \cos t & + & v_3' \sin t & = & 0 \\ 2v_1' e^{2t} & - & v_2' \sin t & + & v_3' \cos t & = & 0 \\ 4v_1' e^{2t} & - & v_2' \cos t & - & v_3' \sin t & = & 5t \end{cases}.$$

Solving this system we get

$$\begin{cases} v_1' &= te^{-2t} \\ v_2' &= t(2\sin t - \cos t) \\ v_3' &= -t(2\cos t + \sin t) \end{cases}.$$

Integrating the equations gives

$$\begin{cases} v_1 &= -\frac{2t+1}{4} e^{-2t} \\ v_2 &= -t(2\cos t + \sin t) - \cos t + 2\sin t \\ v_3 &= t(\cos t - 2\sin t) - 2\cos t - \sin t \end{cases}.$$

Therefore

$$\begin{aligned} y_p(t) &= -\frac{2t+1}{4} e^{-2t} e^{2t} \\ &\quad + (-t(2\cos t + \sin t) - \cos t + 2\sin t)\cos t \\ &\quad + (t(\cos t - 2\sin t) - 2\cos t - \sin t)\sin t \\ &= -\frac{5}{4}(2t+1) \end{aligned}$$

is a particular solution. □

Example 4.7.15. *Find a particular solution of*

$$y''' + y' = \sec^2 t, \quad t \in (-\pi/2, \pi/2).$$

Solution: The roots of the characteristic equation are $r = 0, \pm i$. A set of fundamental solutions is given by

$$\begin{cases} y_1 &= 1 \\ y_2 &= \cos t \\ y_3 &= \sin t \end{cases}.$$

The Wronskian is

$$W(t) = \begin{vmatrix} 1 & \cos t & \sin t \\ 0 & -\sin t & \cos t \\ 0 & -\cos t & -\sin t \end{vmatrix} = 1.$$

Using variation of parameter, let

$$y_p(t) = v_1 + v_2 \cos t + v_3 \sin t,$$

where v_1, v_2, v_3 satisfy

$$\left\{ \begin{array}{rcrcrcc} v_1' & + & v_2' \cos t & + & v_3' \sin t & = & 0 \\ & - & v_2' \sin t & + & v_3' \cos t & = & 0 \\ & - & v_2' \cos t & - & v_3' \sin t & = & \sec^2 t \end{array} \right. .$$

Solving this system we get

$$\left\{ \begin{array}{rcl} v_1' & = & \sec^2 t \\ v_2' & = & -\sec t \\ v_3' & = & -\sec t \tan t \end{array} \right. .$$

Integrating the equations gives

$$\left\{ \begin{array}{rcl} v_1 & = & \tan t \\ v_2 & = & -\ln |\sec t + \tan t| \\ v_3 & = & -\sec t \end{array} \right. .$$

Therefore

$$\begin{aligned} y_p(t) &= \tan t - \cos t \ln |\sec t + \tan t| - \sec t \sin t \\ &= -\cos t \ln |\sec t + \tan t| \end{aligned}$$

is a particular solution. □

Exercise 4.7

1. Write down a suitable form $y_p(t)$ of a particular solution of the following nonhomogeneous second order linear equations.

   (a) $y^{(3)} + y' = 1 - 2\cos t$
   (b) $y^{(3)} - 2y'' + 2y' = t(1 - e^t \cos t)$
   (c) $y^{(4)} - 2y'' + y = te^t$
   (d) $y^{(4)} - 5y'' + 4y = e^t - te^{2t}$
   (e) $y^{(4)} + 2y'' + y = t\cos t$
   (f) $y^{(5)} + 2y^{(3)} + 2y'' = 2t^2$
   (g) $y^{(5)} - y^{(3)} = e^t - 4t^2$

2. Use the method of variation of parameters to find a particular solution of the following nonhomogeneous linear equations.

   (a) $y^{(3)} - y' = t$
   (b) $y^{(3)} - 3y'' + 4y = e^{2t}$
   (c) $y^{(3)} - 2y'' - y' + 2y = e^{4t}$
   (d) $y^{(3)} + y' = \tan x$

5 Eigenvalues and eigenvectors

5.1 Eigenvalues and eigenvectors

Definition 5.1.1 (Eigenvalues and eigenvectors)**.** *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. A number* λ*, which can be a complex number, is called an* **eigenvalue** *of the* $\mathbf{A}$ *if there exists a nonzero vector* $\mathbf{v}$*, which can be a complex vector, such that*

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v},$$

in which case the vector $\mathbf{v}$ *is called an* **eigenvector** *of the matrix* $\mathbf{A}$ *associated with* λ*.*

Example 5.1.2. *Consider*

$$\mathbf{A} = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}.$$

We have

$$\mathbf{A}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Thus $\lambda = 2$ *is an eigenvalue of* $\mathbf{A}$ *and* $(2,1)^T$ *is an eigenvector of* $\mathbf{A}$ *associated with the eigenvalue* $\lambda = 2$*. We also have*

$$\mathbf{A}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Thus $\lambda = 1$ *is an eigenvalue of* $\mathbf{A}$ *and* $(3,2)^T$ *is an eigenvector of* $\mathbf{A}$ *associated with the eigenvalue* $\lambda = 1$*.*

Remarks:

1. An eigenvalue may be zero but an eigenvector is by definition a nonzero vector.
2. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of $\mathbf{A}$ associated with eigenvalue λ, then for any scalars c_1 and c_2, $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ is also an eigenvector of $\mathbf{A}$ associated with eigenvalue λ if it is non-zero.
3. If λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$, then the set of all eigenvectors associated with eigenvalue λ together with the zero vector $\mathbf{0}$ form a vector subspace of $\mathbb{R}^n$. It is called the **eigenspace** of $\mathbf{A}$ associated with eigenvalue λ.

Definition 5.1.3 (Characteristic polynomial)**.** *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. The polynomial function*

$$p(x) = \det(x\mathbf{I} - \mathbf{A})$$

of degree n *is called the* **characteristic polynomial** *of* $\mathbf{A}$*. The* $\det(x\mathbf{I} - \mathbf{A}) = 0$ *is a polynomial equation of degree* n *which is called the* **characteristic equation** *of* $\mathbf{A}$*.*

Theorem 5.1.4. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. The following statements for scalar* λ *are equivalent.*

1. λ *is an eigenvalue of* $\mathbf{A}$*.*
2. *The equation* $(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$ *has nontrivial solution for* $\mathbf{v}$*.*
3. $\text{Null}(\lambda\mathbf{I} - \mathbf{A}) \neq \{\mathbf{0}\}$.
4. *The matrix* $\lambda\mathbf{I} - \mathbf{A}$ *is singular.*
5. λ *is a root of the characteristic equation* $\det(x\mathbf{I} - \mathbf{A}) = 0$

To find the eigenvalues of a square matrix $\mathbf{A}$, we may solve the characteristic equation $\det(x\mathbf{I}-\mathbf{A}) = 0$. For each eigenvalue λ of $\mathbf{A}$, we may find an eigenvector of $\mathbf{A}$ associated with λ by finding a non-trivial solution to $(\lambda\mathbf{I}-\mathbf{A})\mathbf{v} = \mathbf{0}$.

Example 5.1.5. *Find the eigenvalues and associated eigenvectors of the matrix*

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \det(\lambda\mathbf{I}-\mathbf{A}) &= 0 \\ \begin{vmatrix} \lambda-3 & -2 \\ -3 & \lambda+2 \end{vmatrix} &= 0 \\ \lambda^2-\lambda-12 &= 0 \\ \lambda &= 4,-3 \end{aligned}$$

For $\lambda_1 = 4$,

$$\begin{aligned} (4\mathbf{I}-\mathbf{A})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} 1 & -2 \\ -3 & 6 \end{pmatrix}\mathbf{v} &= \mathbf{0} \end{aligned}$$

Thus $\mathbf{v}_1 = (2,1)^T$ is an eigenvector associated with $\lambda_1 = 4$.
For $\lambda_2 = -3$,

$$\begin{aligned} (-3\mathbf{I}-\mathbf{A})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} -6 & -2 \\ -3 & -1 \end{pmatrix}\mathbf{v} &= \mathbf{0} \end{aligned}$$

Thus $\mathbf{v}_2 = (1,-3)^T$ is an eigenvector associated with $\lambda_2 = -3$. □

Example 5.1.6. *Find the eigenvalues and associated eigenvectors of the matrix*

$$\mathbf{A} = \begin{pmatrix} 0 & 8 \\ -2 & 0 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} \lambda & -8 \\ 2 & \lambda \end{vmatrix} &= 0 \\ \lambda^2+16 &= 0 \\ \lambda &= \pm 4i \end{aligned}$$

For $\lambda_1 = 4i$,

$$\begin{aligned} (4i\mathbf{I}-\mathbf{A})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} 4i & -8 \\ 2 & 4i \end{pmatrix}\mathbf{v} &= \mathbf{0} \end{aligned}$$

Thus $\mathbf{v}_1 = (2,i)^T$ is an eigenvector associated with $\lambda_1 = 4i$.
For $\lambda_2 = -4i$,

$$\begin{aligned} (-4i\mathbf{I}-\mathbf{A})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} -4i & -8 \\ 2 & -4i \end{pmatrix}\mathbf{v} &= \mathbf{0} \end{aligned}$$

Thus $\mathbf{v}_2 = (2, -i)^T$ is an eigenvector associated with $\lambda_2 = -4i$. (Note that $\lambda_2 = \bar{\lambda_1}$ and $\mathbf{v}_2 = \bar{\mathbf{v}_1}$ in this example.) □

Remark: For any square matrix $\mathbf{A}$ with real entries, the characteristic polynomial of $\mathbf{A}$ has real coefficients. Thus if $\lambda = \rho + \mu i$, where $\rho, \mu \in \mathbb{R}$, is a complex eigenvalue of $\mathbf{A}$, then its conjugate $\bar{\lambda} = \rho - \mu i$ is also an eigenvalue of $\mathbf{A}$. Furthermore, if $\mathbf{v} = \mathbf{a} + \mathbf{b}i$ is an eigenvector associated with complex eigenvalue λ, then $\bar{\mathbf{v}} = \mathbf{a} - \mathbf{b}i$ is an eigenvector associated with eigenvalue $\bar{\lambda}$.

Example 5.1.7. *Find the eigenvalues and a basis for each eigenspace of the matrix*

$$\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} \lambda - 2 & 3 & -1 \\ -1 & \lambda + 2 & -1 \\ -1 & 3 & \lambda - 2 \end{vmatrix} &= 0 \\ \lambda(\lambda - 1)^2 &= 0 \\ \lambda &= 1, 1, 0 \end{aligned}$$

For $\lambda_1 = \lambda_2 = 1$,

$$\begin{pmatrix} -1 & 3 & -1 \\ -1 & 3 & -1 \\ -1 & 3 & -1 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $\{\mathbf{v}_1 = (3, 1, 0)^T, \mathbf{v}_2 = (-1, 0, 1)^T\}$ constitutes a basis for the eigenspace associated with eigenvalue $\lambda = 1$. For $\lambda_3 = 0$,

$$\begin{pmatrix} -2 & 3 & -1 \\ -1 & 2 & -1 \\ -1 & 3 & -2 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $\{\mathbf{v}_3 = (1, 1, 1)^T\}$ constitutes a basis for the eigenspace associated with eigenvalue $\lambda = 0$.□

In the above example, the characteristic equation of $\mathbf{A}$ has a root 1 of multiplicity two and a root 0 of multiplicity one. For eigenvalue $\lambda = 1$, there associates two linearly independent eigenvectors and for $\lambda = 0$, there associates one linearly independent eigenvector. An important fact in linear algebra is that in general, the number of linearly independent eigenvectors associated with an eigenvalue λ is always less than or equal to the multiplicity of λ as a root of the characteristic equation. A proof of this statement will be given in the next section. However the number of linearly independent eigenvectors associated with an eigenvalue λ can be strictly less than the multiplicity of λ as a root of the characteristic equation as shown by the following two examples.

Example 5.1.8. *Find the eigenvalues and a basis for each eigenspace of the matrix*

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned}\begin{vmatrix} \lambda-2 & -3 \\ 0 & \lambda-2 \end{vmatrix} &= 0 \\ (\lambda-2)^2 &= 0 \\ \lambda &= 2,2\end{aligned}$$

When $\lambda = 2$,

$$\begin{pmatrix} 0 & -3 \\ 0 & 0 \end{pmatrix}\mathbf{v} = \mathbf{0}$$

Thus $\{\mathbf{v} = (1,0)^T\}$ constitutes a basis for the eigenspace associated with eigenvalue $\lambda = 2$. In this example, $\lambda = 2$ is a root of multiplicity two of the characteristic equation but we call only find one linearly independent eigenvector for $\lambda = 2$. □

Example 5.1.9. *Find the eigenvalues and a basis for each eigenspace of the matrix*

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned}\begin{vmatrix} \lambda+1 & -1 & 0 \\ 4 & \lambda-3 & 0 \\ -1 & 0 & \lambda-2 \end{vmatrix} &= 0 \\ (\lambda-2)(\lambda-1)^2 &= 0 \\ \lambda &= 2,1,1\end{aligned}$$

For $\lambda_1 = 2$,

$$\begin{pmatrix} 3 & -1 & 0 \\ 4 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}\mathbf{v} = \mathbf{0}$$

Thus $\{\mathbf{v}_1 = (0,0,1)^T\}$ constitutes a basis for the eigenspace associated with eigenvalue $\lambda = 2$. For $\lambda_2 = \lambda_3 = 1$,

$$\begin{pmatrix} 2 & -1 & 0 \\ 4 & -2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\mathbf{v} = \mathbf{0}$$

Thus $\{\mathbf{v}_2 = (-1,-2,1)^T\}$ constitutes a basis for the eigenspace associated with eigenvalue $\lambda = 1$. Note that here $\lambda = 1$ is a double root but we can only find one linearly independent eigenvector associated with $\lambda = 1$. □

Exercise 5.1

1. Find the eigenvalues and a basis for each eigenspace of each of the following matrices.

(a) $\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$

(b) $\begin{pmatrix} 10 & -9 \\ 6 & -5 \end{pmatrix}$

(c) $\begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix}$

(d) $\begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$

(e) $\begin{pmatrix} 4 & -5 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} -2 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$

(h) $\begin{pmatrix} 3 & 6 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(i) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$

(j) $\begin{pmatrix} -3 & 5 & -5 \\ 3 & -1 & 3 \\ 8 & -8 & 10 \end{pmatrix}$

2. Let $\mathbf{A}$ be a square matrix such that $\mathbf{A}^2 = \mathbf{A}$. Show that if λ is an eigenvalue of $\mathbf{A}$ than $\lambda = 0$ or 1.

3. Let $\mathbf{A}$ be a square matrix.

   (a) Show that if λ is an eigenvalue of $\mathbf{A}$, then λ is also an eigenvalue of $\mathbf{A}^T$.

   (b) Show that $\mathbf{A}$ is non-singular if and only if 0 is not an eigenvalue of $\mathbf{A}$.

   (c) Show that if λ is an eigenvalue of $\mathbf{A}$, then for any non-negative integer k, λ^k is an eigenvalue of $\mathbf{A}^k$.

   (d) Show that if $\mathbf{A}$ is non-singular and λ is an eigenvalue (which is non-zero by (2)) of $\mathbf{A}$, then λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$.

4. Show that if $\mathbf{A}$ is an upper-triangular matrix, then λ is an eigenvalue of $\mathbf{A}$ if and only if λ is equal to one of diagonal entries of $\mathbf{A}$.

5.2 Diagonalization

Definition 5.2.1 (Similar matrices)**.** *Two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be* **similar** *if there exists an invertible (may be complex) matrix $\mathbf{P}$ such that*

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

Theorem 5.2.2. *Similarity of square matrices is an equivalence relation, that is,*

1. *For any square matrix $\mathbf{A}$, we have $\mathbf{A}$ is similar to $\mathbf{A}$;*

2. *If $\mathbf{A}$ is similar to $\mathbf{B}$, then $\mathbf{B}$ is similar to $\mathbf{A}$;*

3. *If $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{B}$ is similar to $\mathbf{C}$, then $\mathbf{A}$ is similar to $\mathbf{C}$.*

Proof.

1. Since $\mathbf{I}$ is a non-singular matrix and $\mathbf{A} = \mathbf{I}^{-1}\mathbf{A}\mathbf{I}$, we have $\mathbf{A}$ is similar to $\mathbf{A}$.

2. If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Now $\mathbf{P}^{-1}$ is a non-singular matrix and $(\mathbf{P}^{-1})^{-1} = \mathbf{P}$. There exists non-singular matrix $\mathbf{P}^{-1}$ such that $(\mathbf{P}^{-1})^{-1}\mathbf{B}\mathbf{P}^{-1} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1} = \mathbf{A}$. Therefore $\mathbf{B}$ is similar to $\mathbf{A}$.

3. If $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{B}$ is similar to $\mathbf{C}$, then there exists non-singular matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ and $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$. Now $\mathbf{P}\mathbf{Q}$ is a non-singular matrix and $(\mathbf{P}\mathbf{Q})^{-1} = \mathbf{Q}^{-1}\mathbf{P}^{-1}$. There exists non-singular matrix $\mathbf{P}\mathbf{Q}$ such that $(\mathbf{P}\mathbf{Q})^{-1}\mathbf{A}(\mathbf{P}\mathbf{Q}) = \mathbf{Q}^{-1}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\mathbf{Q} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} = \mathbf{C}$. Therefore $\mathbf{A}$ is similar to $\mathbf{C}$.

□

Theorem 5.2.3.

1. *The only matrix similar to the zero matrix* $\mathbf{0}$ *is the zero matrix.*
2. *The only matrix similar to the identity matrix* $\mathbf{I}$ *is the identity matrix.*
3. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $a\mathbf{A}$ *is similar to* $a\mathbf{B}$ *for any real number* a*.*
4. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathbf{A}^k$ *is similar to* $\mathbf{B}^k$ *for any non-negative integer* k*.*
5. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are similar non-singular matrices, then* $\mathbf{A}^{-1}$ *is similar to* $\mathbf{B}^{-1}$*.*
6. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathbf{A}^T$ *is similar to* $\mathbf{B}^T$*.*
7. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\det(\mathbf{A}) = \det(\mathbf{B})$*.*
8. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{B})$ *where* $\operatorname{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}$ *is the* **trace***, i.e., the sum of the entries in the diagonal, of* $\mathbf{A}$*.*
9. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathbf{A}$ *and* $\mathbf{B}$ *have the same characteristic equation.*

Proof.

1. Suppose $\mathbf{A}$ is similar to $\mathbf{0}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{0} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Hence $\mathbf{A} = \mathbf{P}\mathbf{0}\mathbf{P}^{-1} = \mathbf{0}$.
2. Similar to (1) and is left as exercise.
3. Exercise.
4. If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. We have

$$\begin{aligned}
\mathbf{B}^k &= \overbrace{(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\cdots(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})}^{k \text{ copies}} \\
&= \mathbf{P}^{-1}\mathbf{A}(\mathbf{P}\mathbf{P}^{-1})\mathbf{A}\mathbf{P}\cdots\mathbf{P}^{-1}\mathbf{A}(\mathbf{P}\mathbf{P}^{-1})\mathbf{A}\mathbf{P} \\
&= \mathbf{P}^{-1}\mathbf{A}\mathbf{I}\mathbf{A}\mathbf{I}\cdots\mathbf{I}\mathbf{A}\mathbf{I}\mathbf{A}\mathbf{P} \\
&= \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}
\end{aligned}$$

   Therefore $\mathbf{A}^k$ is similar to $\mathbf{B}^k$.

5. If $\mathbf{A}$ and $\mathbf{B}$ are similar non-singular matrices, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. We have

$$\begin{aligned}
\mathbf{B}^{-1} &= (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})^{-1} \\
&= \mathbf{P}^{-1}\mathbf{A}^{-1}(\mathbf{P}^{-1})^{-1} \\
&= \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{P}
\end{aligned}$$

   Therefore $\mathbf{A}^{-1}$ is similar to $\mathbf{B}^{-1}$.

6. Exercise.

7. If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Thus $\det(\mathbf{B}) = \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \det(\mathbf{P}^{-1})\det(\mathbf{A})\det(\mathbf{P}) = \det(\mathbf{A})$ since $\det(\mathbf{P}^{-1}) = \det(\mathbf{P})^{-1}$.
8. If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Thus $\mathrm{tr}(\mathbf{B}) = \mathrm{tr}((\mathbf{P}^{-1}\mathbf{A})\mathbf{P}) = \mathrm{tr}(\mathbf{P}(\mathbf{P}^{-1}\mathbf{A})) = \mathrm{tr}(\mathbf{A})$. (Note: It is well-known that $\mathrm{tr}(\mathbf{PQ}) = \mathrm{tr}(\mathbf{QP})$ for any square matrices $\mathbf{P}$ and $\mathbf{Q}$. But in general, it is not always true that $\mathrm{tr}(\mathbf{PQR}) = \mathrm{tr}(\mathbf{QPR})$.)
9. Similar to (7) and is left as exercise.

□

Definition 5.2.4. *An $n \times n$ matrix $\mathbf{A}$ is said to be* **diagonalizable** *if there exists a non-singular (may be complex) matrix $\mathbf{P}$ such that*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

is a diagonal matrix. In this case we say that $\mathbf{P}$ diagonalizes $\mathbf{A}$. In other words, $\mathbf{A}$ is diagonalizable if it is similar to a diagonal matrix.

Theorem 5.2.5. *Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n linearly independent eigenvectors.*

Proof. Let $\mathbf{P}$ be an $n \times n$ matrix and write

$$\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}.$$

where $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are the column vectors of $\mathbf{P}$. First observe that $\mathbf{P}$ is non-singular if and only if $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent (Theorem 3.3.13). Furthermore

$$\begin{aligned}
& \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \text{ is a diagonal matrix.} \\
\Leftrightarrow\quad & \mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D} \text{ for some diagonal matrix } \mathbf{D} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}. \\
\Leftrightarrow\quad & \mathbf{A}\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix}. \\
\Leftrightarrow\quad & \begin{bmatrix} \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix}. \\
\Leftrightarrow\quad & \mathbf{v}_k \text{ is an eigenvector of } \mathbf{A} \text{ associated with eigenvalue } \lambda_k \text{ for } k = 1, 2, \cdots, n.
\end{aligned}$$

Therefore $\mathbf{P}$ diagonalizes $\mathbf{A}$ if and only if $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent eigenvectors of $\mathbf{A}$. □

Example 5.2.6. *Diagonalize the matrix*

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}.$$

Solution: We have seen in Example 5.1.5 that $\mathbf{A}$ has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = -3$ associated with linearly independent eigenvectors $\mathbf{v}_1 = (2, 1)^T$ and $\mathbf{v}_2 = (1, -3)^T$ respectively. Thus the matrix

$$\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix}. \end{aligned}$$

□

Example 5.2.7. *Diagonalize the matrix*

$$\mathbf{A} = \begin{pmatrix} 0 & 8 \\ -2 & 0 \end{pmatrix}.$$

Solution: We have seen in Example 5.1.6 that $\mathbf{A}$ has eigenvalues $\lambda_1 = 4i$ and $\lambda_2 = -4i$ associated with linearly independent eigenvectors $\mathbf{v}_1 = (2, i)^T$ and $\mathbf{v}_2 = (2, -i)^T$ respectively. Thus the matrix

$$\mathbf{P} = \begin{pmatrix} 2 & 2 \\ i & -i \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 4i & 0 \\ 0 & -4i \end{pmatrix}.$$

□

Example 5.2.8. *We have seen in Example 5.1.7 that*

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

has only one linearly independent eigenvector. Therefore it is not diagonalizable.

Example 5.2.9. *Diagonalize the matrix*

$$\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}.$$

Solution: We have seen in Example 5.1.6 that $\mathbf{A}$ has eigenvalues $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0$. For $\lambda_1 = \lambda_2 = 1$, there are two linearly independent eigenvectors $\mathbf{v}_1 = (3, 1, 0)^T$ and $\mathbf{v}_2 = (-1, 0, 1)^T$. For $\lambda_3 = 0$, there associated one linearly independent eigenvector $\mathbf{v}_3 = (1, 1, 1)^T$. The three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent eigenvectors. Thus the matrix

$$\mathbf{P} = \begin{pmatrix} 3 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

Example 5.2.10. *Show that the matrix*

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

is not diagonalizable.

Solution: One can show that there is at most one linearly independent eigenvector. Alternatively, one can argue in the following way. The characteristic equation of $\mathbf{A}$ is $(r-1)^2 = 0$. Thus $\lambda = 1$ is the only eigenvalue of $\mathbf{A}$. Hence if $\mathbf{A}$ is diagonalizable by $\mathbf{P}$, then $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{I}$. But then $\mathbf{A} = \mathbf{P}\mathbf{I}\mathbf{P}^{-1} = \mathbf{I}$ which leads to a contradiction. □

Theorem 5.2.11. *Suppose that eigenvectors* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *are associated with the distinct eigenvalues* $\lambda_1, \lambda_2, \cdots, \lambda_k$ *of a matrix* $\mathbf{A}$. *Then* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *are linearly independent.*

Proof. We prove the theorem by induction on k. The theorem is obviously true when $k = 1$. Now assume that the theorem is true for any set of $k-1$ eigenvectors. Suppose

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

Multiplying $\mathbf{A} - \lambda_k\mathbf{I}$ to the left on both sides, we have

$$\begin{aligned} c_1(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_1 + c_2(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_2 + \cdots + c_{k-1}(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_{k-1} + c_k(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_k &= \mathbf{0} \\ c_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + c_2(\lambda_2 - \lambda_k)\mathbf{v}_2 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1} &= \mathbf{0} \end{aligned}$$

Note that $(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_k = \mathbf{0}$ since $\mathbf{v}_k$ is an eigenvector associated with λ_k. From the induction hypothesis, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{k-1}$ are linearly independent. Thus

$$c_1(\lambda_1 - \lambda_k) = c_2(\lambda_2 - \lambda_k) = \cdots = c_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Since $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct, $\lambda_1 - \lambda_k, \lambda_2 - \lambda_k, \cdots, \lambda_{k-1} - \lambda_k$ are all nonzero. Hence

$$c_1 = c_2 = \cdots = c_{k-1} = 0.$$

It follows then that c_k is also equal to zero because $\mathbf{v}_k$ is a nonzero vector. Therefore $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent. □

The above theorem gives a sufficient condition for a matrix to be diagonalizable.

Theorem 5.2.12. *If the* $n \times n$ *matrix* $\mathbf{A}$ *has* n *distinct eigenvalues, then it is diagonalizable.*

Definition 5.2.13 (Algebraic and geometric multiplicity)**.** *Let* $\mathbf{A}$ *be a square matrix and* λ *be an eigenvalue of* $\mathbf{A}$, *in other words,* λ *is a root of the characteristic equation of* $\mathbf{A}$.

1. *The* **algebraic multiplicity** *of* λ *is the multiplicity of* λ *being a root of the characteristic equation of* $\mathbf{A}$. *The algebraic multiplicity of* λ *is denoted by* $m_a(\lambda)$.

2. *The* **geometric multiplicity** *of λ is the dimension of the eigenspace associated to eigenvalue λ, that is, the maximum number of linearly independent eigenvectors associated with eigenvalue λ. The geometric multiplicity of λ is denoted by $m_g(\lambda)$.*

We have the following important theorem concerning the algebraic multiplicity and geometric multiplicity of an eigenvalue.

Theorem 5.2.14. *Let $\mathbf{A}$ be an $n \times n$ matrix and λ be an eigenvalue of $\mathbf{A}$. Then we have*

$$1 \leq m_g(\lambda) \leq m_a(\lambda)$$

where $m_g(\lambda)$ and $m_a(\lambda)$ are the geometric and algebraic multiplicity of λ respectively. In other words, the maximum number of linearly independent eigenvectors associated with λ is less than or equal to the algebraic multiplicity of λ as a root of the characteristic equation of $\mathbf{A}$.

Proof. Suppose there are k linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k \in \mathbb{R}^n$ of $\mathbf{A}$ associated with λ. We are going to prove that the algebraic multiplicity of λ is at least k. Let $\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \cdots, \mathbf{u}_n \in \mathbb{R}^n$ be vectors such that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k, \mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \cdots, \mathbf{u}_n$ constitute a basis for $\mathbb{R}^n$. Using these vectors as column vectors, the $n \times n$ matrix

$$\mathbf{P} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k, \mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \cdots, \mathbf{u}_n]$$

is non-singular (Theorem 3.3.13). Consider the matrix $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ which must be of the form

$$\mathbf{B} = \begin{pmatrix} \lambda\mathbf{I} & \mathbf{C} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}$$

where $\mathbf{I}$ is the $k \times k$ identity matrix, $\mathbf{0}$ is the $(n-k) \times k$ zero matrix, $\mathbf{C}$ is a $k \times (n-k)$ matrix and $\mathbf{D}$ is an $(n-k) \times (n-k)$ matrix. Note that since $\mathbf{A}$ and $\mathbf{B}$ are similar, the characteristic equation of $\mathbf{A}$ and $\mathbf{B}$ are the same (Theorem 5.2.3). Observe that

$$\det(x\mathbf{I} - \mathbf{B}) = \begin{vmatrix} (x-\lambda)\mathbf{I} & \mathbf{C} \\ \mathbf{0} & x\mathbf{I} - \mathbf{D} \end{vmatrix} = (x-\lambda)^k \det(x\mathbf{I} - \mathbf{D})$$

We see that the algebraic multiplicity of λ as root of the characteristic equation of $\mathbf{B}$ is as least k and therefore the algebraic multiplicity of λ as root of the characteristic equation of $\mathbf{A}$ is as least k. □

If the geometric multiplicity is equal to the algebraic multiplicity for each eigenvalue of $\mathbf{A}$, then $\mathbf{A}$ is diagonalizable.

Theorem 5.2.15. *Let $\mathbf{A}$ be an $n \times n$ matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ be the distinct roots of the characteristic equation of $\mathbf{A}$ of multiplicity $n_1, n_2, \cdots, n_k$ respectively. In other words, the characteristic polynomial of $\mathbf{A}$ is*

$$(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_k)^{n_k}$$

Then $\mathbf{A}$ is diagonalizable if and only if for each $1 \leq i \leq k$, there exists n_i linearly independent eigenvectors associated with eigenvalue λ_i.

Proof. Suppose for each eigenvalue λ_i, there exists n_i linearly independent eigenvectors associated with eigenvalue λ_i. Putting all these eigenvectors together, we obtain a set of n eigenvectors of $\mathbf{A}$. Using the same argument in the proof of Theorem 5.2.11, one may prove that these n eigenvectors are linearly independent. Therefore $\mathbf{A}$ is diagonalizable (Theorem 5.2.5).

Suppose $\mathbf{A}$ is diagonalizable. Then there exists n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ of $\mathbf{A}$ (Theorem 5.2.5). Among these vectors, at most n_i of them are associated with λ_i for each $1 \leq i \leq k$. Since $n_1 + n_2 + \cdots + n_k = n$, we must have exactly n_i of them associated with λ_i for each i. Therefore there are n_i linearly independent eigenvectors associated with λ_i for each i. □

We conclude this section by giving three more theorems without proof.

Theorem 5.2.16. *All eigenvalues of a symmetric matrix are real.*

Theorem 5.2.17. *Any symmetric matrix is diagonalizable (by orthogonal matrix).*

Exercise 5.2

1. Diagonalize the following matrices.

   (a) $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$
   (b) $\begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$
   (c) $\begin{pmatrix} 5 & -4 \\ 2 & -1 \end{pmatrix}$
   (d) $\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 6 & 11 & 6 \end{pmatrix}$
   (e) $\begin{pmatrix} 3 & -2 & 0 \\ 0 & 1 & 0 \\ -4 & 4 & 1 \end{pmatrix}$
   (f) $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$
   (g) $\begin{pmatrix} 7 & -8 & 3 \\ 6 & -7 & 3 \\ 2 & -2 & 2 \end{pmatrix}$

2. Show that that following matrices are not diagonalizable.

   (a) $\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$
   (b) $\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$
   (c) $\begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix}$

3. Let $\mathbf{A}$ and $\mathbf{B}$ be non-singular matrices. Prove that if $\mathbf{A}$ is similar to $\mathbf{B}$, then $\mathbf{A}^{-1}$ is similar to $\mathbf{B}$.

4. Suppose $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{C}$ is similar to $\mathbf{D}$. Explain whether it is always true that $\mathbf{AC}$ is similar to $\mathbf{BD}$.

5. Suppose $\mathbf{A}$ and $\mathbf{B}$ are similar matrices. Show that if λ is an eigenvalue of $\mathbf{A}$, then λ is an eigenvalue of $\mathbf{B}$.

6. Let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ matrices. Show that $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$. (Note: In general, it is not always true that $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BAC})$.)

7. Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a 2×2 matrix. Show that if $(a-d)^2 + 4bc \neq 0$, then $\mathbf{A}$ is diagonalizable.

8. Suppose $\mathbf{P}$ diagonalizes two matrices $\mathbf{A}$ and $\mathbf{B}$ simultaneously. Prove that $\mathbf{AB} = \mathbf{BA}$.

9. A square matrix $\mathbf{A}$ is said to be nilpotent if there exists positive integer k such that $\mathbf{A}^k = \mathbf{0}$. Prove that any non-zero nilpotent matrix is not diagonalizable.

10. Prove that if $\mathbf{A}$ is a non-singular matrix, then for any matrix $\mathbf{B}$, we have $\mathbf{AB}$ is similar to $\mathbf{BA}$.

11. Show that there exists matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{AB}$ is not similar to $\mathbf{BA}$.

12. Show that $\mathbf{AB}$ and $\mathbf{BA}$ have the same characteristic equation.

5.3 Power of matrices

Let $\mathbf{A}$ be an $n \times n$ matrix and $\mathbf{P}$ be a matrix diagonalizes $\mathbf{A}$, i.e.,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

is a diagonal matrix. Then

$$\mathbf{A}^k = \left(\mathbf{P}\mathbf{D}\mathbf{P}^{-1}\right)^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}.$$

Example 5.3.1. *Find* $\mathbf{A}^5$ *if*

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}.$$

Solution: From Example 5.2.6,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix}$$

where

$$\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}.$$

Thus

$$\begin{aligned} \mathbf{A}^5 &= \mathbf{P}\mathbf{D}^5\mathbf{P}^{-1} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix}^5\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} 1024 & 0 \\ 0 & -243 \end{pmatrix}\frac{1}{-7}\begin{pmatrix} -3 & -1 \\ -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 843 & 362 \\ 543 & -62 \end{pmatrix} \end{aligned}$$

□

Example 5.3.2. *Find* $\mathbf{A}^5$ *if*

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{pmatrix}.$$

Solution: Diagonalizing $\mathbf{A}$, we have

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

where

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Thus

$$\begin{aligned}\mathbf{A}^5 &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^5 \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 243 & 0 & 0 \\ 0 & 32 & 0 \\ 0 & 0 & 32 \end{pmatrix} \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ -1 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 454 & -422 & 211 \\ 422 & -390 & 211 \\ 422 & -422 & 243 \end{pmatrix}\end{aligned}$$

□

Example 5.3.3. *Consider a metropolitan area with a constant total population of 1 million individuals. This area consists of a city and its suburbs, and we want to analyze the changing urban and suburban populations. Let C_k denote the city population and S_k the suburban population after k years. Suppose that each year 15% of the people in the city move to the suburbs, whereas 10% of the people in the suburbs move to the city. Then it follows that*

$$\begin{cases} C_{k+1} &= 0.85C_k + 0.1S_k \\ S_{k+1} &= 0.15C_k + 0.9S_k \end{cases}$$

Find the urban and suburban populations after a long time.

Solution: Let $\mathbf{x}_k = (C_k, S_k)^T$ be the population vector after k years. Then

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}^2\mathbf{x}_{k-2} = \cdots = \mathbf{A}^k\mathbf{x}_0,$$

where

$$\mathbf{A} = \begin{pmatrix} 0.85 & 0.1 \\ 0.15 & 0.9 \end{pmatrix}.$$

Solving the characteristic equation, we have

$$\begin{aligned}\begin{vmatrix} \lambda - 0.85 & -0.1 \\ -0.15 & \lambda - 0.9 \end{vmatrix} &= 0 \\ \lambda^2 - 1.75\lambda + 0.75 &= 0 \\ \lambda &= 1, 0.75\end{aligned}$$

Hence the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 0.75$. By solving $\mathbf{A} - \lambda\mathbf{I} = \mathbf{0}$, the associated eigenvectors are $\mathbf{v}_1 = (2,3)^T$ and $\mathbf{v}_2 = (-1,1)^T$ respectively. Thus

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0.75 \end{pmatrix},$$

where

$$\mathbf{P} = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}.$$

When k is very large

$$\begin{aligned}
\mathbf{A}^k &= \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} \\
&= \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.75 \end{pmatrix}^k \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.75^k \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix} \\
&\simeq \frac{1}{5} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{A}^k\mathbf{x}_0 \\
&\simeq \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} C_0 \\ S_0 \end{pmatrix} \\
&= (C_0 + S_0) \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} \\
&= \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}
\end{aligned}$$

That mean whatever the initial distribution of population is, the long-term distribution consists of 40% in the city and 60% in the suburbs. □

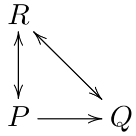
An $n \times n$ matrix is called a **stochastic matrix** if it has nonnegative entries and the sum of the elements in each column is one. A Markov process is a stochastic process having the property that given the present state, future states are independent of the past states. A Markov process can be described by a **Morkov chain** which consists of a sequence of vectors $\mathbf{x}_k$, $k = 0, 1, 2, \cdots$, satisfying

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}^2\mathbf{x}_{k-2} = \cdots = \mathbf{A}^k\mathbf{x}_0,$$

for some stochastic matrix $\mathbf{A}$. The vector $\mathbf{x}_k$ is called the **state vector** and $\mathbf{A}$ is called the **transition matrix**.

The PageRank algorithm was developed by Larry Page and Sergey Brin and is used to rank the web sites in the Google search engine. It is a probability distribution used to represent the likelihood that a person randomly clicking links will arrive at any particular page. The linkage of the web sites in the web can be represented by a linkage matrix $\mathbf{A}$. The probability distribution of a person to arrive the web sites are given by $\mathbf{A}^k\mathbf{x}_0$ for a sufficiently large k and is independent of the initial distribution $\mathbf{x}_0$.

Example 5.3.4 (PageRank)**.** *Consider a small web consisting of three pages P, Q and R, where page P links to the pages Q and R, page Q links to page R and page R links to page P and Q. Assume that a person has a probability of 0.5 to stay on each page and the probability of going to other pages are evenly distributed to the pages which are linked to it. Find the page rank of the three web pages.*



Solution: Let p_k, q_k and r_k be the number of people arrive the web pages P, Q and R respectively after k iteration. Then

$$\begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} p_{k-1} \\ q_{k-1} \\ r_{k-1} \end{pmatrix}.$$

Thus

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \cdots = \mathbf{A}^k \mathbf{x}_0,$$

where

$$\mathbf{A} = \begin{pmatrix} 0.5 & 0 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.5 & 0.5 \end{pmatrix}$$

is the linkage matrix and

$$\mathbf{x}_0 = \begin{pmatrix} p_0 \\ q_0 \\ r_0 \end{pmatrix}$$

is the initial state. Solving the characteristic equation of $\mathbf{A}$, we have

$$\begin{aligned} \begin{vmatrix} \lambda - 0.5 & 0 & -0.25 \\ -0.25 & \lambda - 0.5 & -0.25 \\ -0.25 & -0.5 & \lambda - 0.5 \end{vmatrix} &= 0 \\ \lambda^3 - 1.5\lambda^2 + 0.5625\lambda - 0.0625 &= 0 \\ (\lambda - 1)(\lambda - 0.25)^2 &= 0 \\ \lambda &= 1 \text{ or } 0.25 \end{aligned}$$

For $\lambda_1 = 1$, we solve

$$\begin{pmatrix} -0.5 & 0 & 0.25 \\ 0.25 & -0.5 & 0.25 \\ 0.25 & 0.5 & -0.5 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

and $\mathbf{v}_1 = (2, 3, 4)^T$ is an eigenvector of $\mathbf{A}$ associated with $\lambda_1 = 1$.
For $\lambda_2 = \lambda_3 = 0.25$, we solve

$$\begin{pmatrix} 0.25 & 0 & 0.25 \\ 0.25 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

and there is only one linearly independent eigenvector $\mathbf{v}_2 = (1, 0, -1)^T$ associated with 0.25. Thus $\mathbf{A}$ is not diagonalizable. However we may take

$$\mathbf{P} = \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & -4 \\ 4 & -1 & 0 \end{pmatrix}$$

and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.25 & 1 \\ 0 & 0 & 0.25 \end{pmatrix}.$$

(**J** is called the Jordan normal form of **A**.) When k is sufficiently large, we have

$$\begin{aligned}\mathbf{A}^k &= \mathbf{P}\mathbf{J}^k\mathbf{P}^{-1}\\ &= \begin{pmatrix} 2 & 1 & 4\\ 3 & 0 & -4\\ 4 & -1 & 0\end{pmatrix}\begin{pmatrix} 1 & 0 & 0\\ 0 & 0.25 & 1\\ 0 & 0 & 0.25\end{pmatrix}^k\begin{pmatrix} 2 & 1 & 4\\ 3 & 0 & -4\\ 4 & -1 & 0\end{pmatrix}^{-1}\\ &\simeq \begin{pmatrix} 2 & 1 & 4\\ 3 & 0 & -4\\ 4 & -1 & 0\end{pmatrix}\begin{pmatrix} 1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1/9 & 1/9 & 1/9\\ 4/9 & 4/9 & -5/9\\ 1/12 & -1/6 & 1/12\end{pmatrix}\\ &= \begin{pmatrix} 2/9 & 2/9 & 2/9\\ 3/9 & 3/9 & 3/9\\ 4/9 & 4/9 & 4/9\end{pmatrix}\end{aligned}$$

Thus after sufficiently many iteration, the number of people arrive the web pages are given by

$$\mathbf{A}^k\mathbf{x}_0 \simeq \begin{pmatrix} 2/9 & 2/9 & 2/9\\ 3/9 & 3/9 & 3/9\\ 4/9 & 4/9 & 4/9\end{pmatrix}\begin{pmatrix} p_0\\ q_0\\ r_0\end{pmatrix} = (p_0+q_0+r_0)\begin{pmatrix} 2/9\\ 3/9\\ 4/9\end{pmatrix}.$$

Note that the ratio does not depend on the initial state. The PageRank of the web pages P, Q and R are 2/9, 3/9 and 4/9 respectively. ☐

Example 5.3.5 (Fibonacci sequence)**.** *The Fibonacci sequence is defined by the recurrence relation*

$$\begin{cases} F_{k+2} = F_{k+1} + F_k, & \text{for } k \geq 0\\ F_0 = 0, F_1 = 1.\end{cases}$$

Find the general term of the Fibonacci sequence.

Solution: The recurrence relation can be written as

$$\begin{pmatrix} F_{k+2}\\ F_{k+1}\end{pmatrix} = \begin{pmatrix} 1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} F_{k+1}\\ F_k\end{pmatrix}.$$

for $k \geq 0$. If we let

$$\mathbf{A} = \begin{pmatrix} 1 & 1\\ 1 & 0\end{pmatrix} \text{ and } \mathbf{x}_k = \begin{pmatrix} F_{k+1}\\ F_k\end{pmatrix},$$

then

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, & \text{for } k \geq 0\\ \mathbf{x}_0 = \begin{pmatrix} 1\\ 0\end{pmatrix}.\end{cases}$$

It follows that

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}^2\mathbf{x}_{k-2} = \cdots = \mathbf{A}^k\mathbf{x}_0.$$

To find $\mathbf{A}^k$, we diagonalize $\mathbf{A}$ and obtain

$$\begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2}\\ 1 & 1\end{pmatrix}^{-1}\begin{pmatrix} 1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2}\\ 1 & 1\end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0\\ 0 & \frac{1-\sqrt{5}}{2}\end{pmatrix}$$

Hence

$$\begin{aligned}\mathbf{A}^k &= \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} & \left(\frac{1+\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-1} \end{pmatrix}\end{aligned}$$

Now

$$\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0 = \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} \end{pmatrix},$$

we have

$$F_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k \right).$$

□

Note that we have

$$\lim_{k\to\infty} \frac{F_{k+1}}{F_k} = \lim_{k\to\infty} \frac{((1+\sqrt{5})/2)^{k+1} - ((1-\sqrt{5})/2)^{k+1}}{((1+\sqrt{5})/2)^{k} - ((1-\sqrt{5})/2)^{k}} = \frac{1+\sqrt{5}}{2}$$

which links the Fibonacci sequence with the number $\frac{1+\sqrt{5}}{2} \approx 1.61803$ which is called the golden ratio.

Exercise 5.3

1. Compute $\mathbf{A}^5$ where $\mathbf{A}$ is the given matrix.
   (a) $\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$
   (b) $\begin{pmatrix} 6 & -6 \\ 4 & -4 \end{pmatrix}$
   (c) $\begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix}$
   (d) $\begin{pmatrix} 1 & -5 \\ 1 & -1 \end{pmatrix}$
   (e) $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 3 & 6 & -3 \end{pmatrix}$
   (f) $\begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & -2 & 2 \end{pmatrix}$
   (g) $\begin{pmatrix} 4 & -3 & 1 \\ 2 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

2. Suppose $\mathbf{A}$ is a stochastic matrix, that is $\mathbf{A}$ is a square matric with non-negative entries and the sum of entries in each column is one.
   (a) Prove that $\lambda = 1$ is an eigenvalue of $\mathbf{A}$.
   (b) Prove that if all entries of $\mathbf{A}$ are positive, then the eigenspace associated with $\lambda = 1$ is of dimension 1.

5.4 Cayley-Hamilton theorem

Let $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial. We may consider $p(x)$ as a matrix valued function with square matrix input and write $p(\mathbf{A}) = a_n \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1 \mathbf{A} + a_0$ for square matrix $\mathbf{A}$.

Definition 5.4.1. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. The* **minimal polynomial** *of* $\mathbf{A}$ *is a nonzero polynomial* $m(x)$ *of minimum degree with leading coefficient 1 satisfying* $m(\mathbf{A}) = \mathbf{0}$*.*

One may ask whether there always exists a nonzero polynomial $p(x)$ with $p(\mathbf{A}) = \mathbf{0}$ for every square matrix $\mathbf{A}$. This is true because we have the Cayley-Hamilton theorem which is one of the most notable theorems in linear algebra.

Theorem 5.4.2 (Cayley-Hamilton theorem)**.** *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix and* $p(x) = \det(x\mathbf{I} - \mathbf{A})$ *be its characteristic polynomial. Then* $p(\mathbf{A}) = \mathbf{0}$*. Moreover, we have* $m(x)$ *divides* $p(x)$ *where* $m(x)$ *is the minimal polynomial of* $\mathbf{A}$*.*

Proof. Let $\mathbf{B} = x\mathbf{I} - \mathbf{A}$ and

$$p(x) = \det(\mathbf{B}) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

be the characteristic polynomial of $\mathbf{A}$. Consider $\mathbf{B}$ as an $n \times n$ matrix whose entries are polynomial in x. Then the adjoint $\mathrm{adj}(\mathbf{B})$ of $\mathbf{B}$ is an $n \times n$ matrix with polynomials of degree at most $n-1$ as entries. We may also consider $\mathrm{adj}(\mathbf{B})$ as a polynomial of degree $n-1$ in x with matrix coefficients

$$\mathrm{adj}(\mathbf{B}) = \mathbf{B}_{n-1}x^{n-1} + \cdots + \mathbf{B}_1x + \mathbf{B}_0$$

where the coefficients $\mathbf{B}_i$ are $n \times n$ constant matrices. On one hand, we have

$$\begin{aligned} \det(\mathbf{B})\mathbf{I} &= (x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0)\mathbf{I} \\ &= \mathbf{I}x^n + a_{n-1}\mathbf{I}x^{n-1} + \cdots + a_1\mathbf{I}x + a_0\mathbf{I} \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \mathbf{B}\mathrm{adj}(\mathbf{B}) &= (x\mathbf{I} - \mathbf{A})(\mathbf{B}_{n-1}x^{n-1} + \cdots + \mathbf{B}_1x + \mathbf{B}_0) \\ &= \mathbf{B}_{n-1}x^n + (\mathbf{B}_{n-2} - \mathbf{A}\mathbf{B}_{n-1})x^{n-1} + \cdots + (\mathbf{B}_0 - \mathbf{A}\mathbf{B}_1)x - \mathbf{A}\mathbf{B}_0 \end{aligned}$$

By Theorem 2.4.18, we have

$$\det(\mathbf{B})\mathbf{I} = \mathbf{B}\mathrm{adj}(\mathbf{B})$$

By comparing the coefficients of the above equality, we get

$$\begin{aligned} \mathbf{I} &= \mathbf{B}_{n-1} \\ a_{n-1}\mathbf{I} &= \mathbf{B}_{n-2} - \mathbf{A}\mathbf{B}_{n-1} \\ &\vdots \\ a_1\mathbf{I} &= \mathbf{B}_0 - \mathbf{A}\mathbf{B}_1 \\ a_0\mathbf{I} &= -\mathbf{A}\mathbf{B}_0 \end{aligned}$$

If we multiply the first equation by $\mathbf{A}^n$, the second by $\mathbf{A}^{n-1}$, and so on, and the last one by $\mathbf{I}$, and then add up the resulting equations, we obtain

$$\begin{aligned} p(\mathbf{A}) &= \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{I} \\ &= \mathbf{A}^n\mathbf{B}_{n-1} + (\mathbf{A}^{n-1}\mathbf{B}_{n-2} - \mathbf{A}^n\mathbf{B}_{n-1}) + \cdots + (\mathbf{A}\mathbf{B}_0 - \mathbf{A}^2\mathbf{B}_1) - \mathbf{A}\mathbf{B}_0 \\ &= \mathbf{0}. \end{aligned}$$

For the second statement, by division algorithm there exists polynomials $q(x)$ and $r(x)$ such that

$$p(x) = q(x)m(x) + r(x)$$

with $r(x) \equiv 0$ or $\deg(r(x)) < \deg(m(x))$. Suppose $r(x)$ is a non-zero polynomial. Let k be a non-zero constant such that $kr(x)$ has leading coefficient 1. Now $kr(\mathbf{A}) = kp(\mathbf{A}) - kq(\mathbf{A})m(\mathbf{A}) = \mathbf{0}$. This contradicts the minimality of $m(x)$. Hence $r(x) \equiv 0$ which implies that $m(x)$ divides $p(x)$. ☐

Let $\mathbf{A}$ be an $n \times n$ matrix and $p(x)$ be the characteristic polynomial of $\mathbf{A}$. Since the eigenvalues of $\mathbf{A}$ are exactly the zeros of its characteristic polynomial (Theorem 5.1.4), we have $p(x) = (x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_k)^{n_k}$ where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the eigenvalues of $\mathbf{A}$. The Cayley-Hamilton theorem (Theorem 5.4.2) gives us a way to find the minimal polynomial from the characteristic polynomial.

Theorem 5.4.3. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix and*

$$p(x) = (x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_k)^{n_k}$$

be the characteristic polynomial of $\mathbf{A}$*, where* $\lambda_1, \lambda_2, \cdots, \lambda_k$ *are distinct eigenvalues of* $\mathbf{A}$*. Then the minimal polynomial of* $\mathbf{A}$ *is of the form*

$$m(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$$

where $1 \leq m_i \leq n_i$ *for any* $i = 1, 2, \cdots, k$.

Proof. By Cayley-Hamilton theorem (Theorem 5.4.2), the minimal polynomial $m(x)$ divides the characteristic polynomial $p(x)$ and thus

$$m(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$$

where $m_i \leq n_i$ for any $i = 1, 2, \cdots, k$. It remains to prove $m_i \neq 0$ for any i. Now suppose $m_i = 0$ for some i. Since λ_i is an eigenvalue of $\mathbf{A}$, there exists eigenvector $\mathbf{v}_i$ such that $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$. Then

$$\begin{aligned} m(\mathbf{A}) &= \mathbf{0} \\ m(\mathbf{A})\mathbf{v}_i &= \mathbf{0} \\ (\mathbf{A}-\lambda_1\mathbf{I})^{m_1}\cdots(\mathbf{A}-\lambda_{i-1}\mathbf{I})^{m_{i-1}}(\mathbf{A}-\lambda_{i+1}\mathbf{I})^{m_{i+1}}\cdots(\mathbf{A}-\lambda_k\mathbf{I})^{m_k}\mathbf{v}_i &= \mathbf{0} \\ (\lambda_i-\lambda_1)^{m_1}\cdots(\lambda_i-\lambda_{i-1})^{m_{i-1}}(\lambda_i-\lambda_{i+1})^{m_{i+1}}\cdots(\lambda_i-\lambda_k)^{m_k}\mathbf{v}_i &= \mathbf{0} \end{aligned}$$

which is a contradiction since $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct and $\mathbf{v}_i$ is nonzero. Therefore we proved that $m_i \geq 1$ for any $i = 1, 2, \cdots, k$ and the proof of the theorem is complete. ☐

Example 5.4.4. *Let*

$$\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}$$

Find the minimal polynomial of $\mathbf{A}$.

Solution: The characteristic polynomial of $\mathbf{A}$ is

$$p(x) = x(x-1)^2.$$

Thus the minimal polynomial is either

$$x(x-1) \quad \text{or} \quad x(x-1)^2.$$

By direct computation

$$\mathbf{A}(\mathbf{A}-\mathbf{I}) = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 \\ 1 & -3 & 1 \\ 1 & -3 & 1 \end{pmatrix} = \mathbf{0}.$$

Hence the minimal polynomial of $\mathbf{A}$ is $m(x) = x(x-1)$. □

Minimal polynomial can be used to characterize diagonalizable matrices.

Theorem 5.4.5. *Let* $\mathbf{A}$ *be a square matrix and* $\lambda_1, \lambda_2, \cdots, \lambda_k$ *be distinct eigenvalues of* $\mathbf{A}$. *Let*

$$m(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$$

be the minimal polynomial of $\mathbf{A}$. *Then* $\mathbf{A}$ *is diagonlizable if and only if* $m_i = 1$ *for any* $i = 1, 2, \cdots, k$. *In other words, a square matrix is diagonlizable if and only if its minimal polynomial is a product of distinct linear factors.*

Proof. Suppose $\mathbf{A}$ is an $n \times n$ matrix which is diagonalizable. Then there exists (Theorem 5.2.5) n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ of $\mathbf{A}$ in $\mathbb{R}^n$. Now for each $j = 1, 2, \cdots, n$, we have $\mathbf{A}\mathbf{v}_j = \lambda_i \mathbf{v}_j$ for some i and hence

$$(\mathbf{A}-\lambda_1\mathbf{I})\cdots(\mathbf{A}-\lambda_i\mathbf{I})\cdots(\mathbf{A}-\lambda_k\mathbf{I})\mathbf{v}_j = \mathbf{0}$$

Since $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent, they constitute a basis (Theorem 3.4.7) for $\mathbf{R}^n$. Thus any vector in $\mathbb{R}^n$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ which implies that

$$(\mathbf{A}-\lambda_1\mathbf{I})(\mathbf{A}-\lambda_2\mathbf{I})\cdots(\mathbf{A}-\lambda_k\mathbf{I})\mathbf{v} = \mathbf{0}$$

for any $\mathbf{v} \in \mathbb{R}^n$. It follows that we must have

$$(\mathbf{A}-\lambda_1\mathbf{I})(\mathbf{A}-\lambda_2\mathbf{I})\cdots(\mathbf{A}-\lambda_k\mathbf{I}) = \mathbf{0}$$

(Note that by Theorem 5.4.3 we always have $m_i \geq 1$ for each i.) Therefore the minimal polynomial of $\mathbf{A}$ is

$$m(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k)$$

On the other hand, suppose the minimal polynomial of $\mathbf{A}$ is

$$m(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k)$$

Then

$$(\mathbf{A}-\lambda_1\mathbf{I})(\mathbf{A}-\lambda_2\mathbf{I})\cdots(\mathbf{A}-\lambda_k\mathbf{I}) = \mathbf{0}$$

For each $i = 1, 2, \cdots, k$, let $n_i = \text{nullity}(\mathbf{A}-\lambda_i\mathbf{I})$ be the nullity of $\mathbf{A}-\lambda_i\mathbf{I}$ and let $\{\mathbf{v}_{i1}, \mathbf{v}_{i2}, \cdots, \mathbf{v}_{in_i}\}$ be a basis for the eigenspace $\text{Null}(\mathbf{A}-\lambda_i\mathbf{I})$ of $\mathbf{A}$ associated with eigenvalue λ_i. Then using the same argument in the proof of Theorem 5.2.11, we have

$$\mathbf{v}_{11}, \cdots, \mathbf{v}_{1n_1}, \mathbf{v}_{21}, \cdots, \mathbf{v}_{2n_2}, \cdots, \mathbf{v}_{k1}, \cdots, \mathbf{v}_{kn_k}$$

are linearly independent vectors in $\mathbb{R}^n$. This implies that

$$n_1 + n_2 + \cdots + n_k \leq n$$

Moreover by Theorem 3.5.10, we have

$$n = \text{nullity}((\mathbf{A}-\lambda_1\mathbf{I})(\mathbf{A}-\lambda_2\mathbf{I})\cdots(\mathbf{A}-\lambda_k\mathbf{I})) \leq n_1 + n_2 + \cdots + n_k$$

Combining the two inequalities, we have

$$n_1 + n_2 + \cdots + n_k = n$$

Hence there exists n linearly independent eigenvectors of $\mathbf{A}$. Therefore $\mathbf{A}$ is diagonalizable by Theorem 5.2.5. □

Example 5.4.6. *Let*

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{pmatrix}$$

Find the minimal polynomial of $\mathbf{A}$*. Then express* $\mathbf{A}^4$ *and* $\mathbf{A}^{-1}$ *as a polynomial in* $\mathbf{A}$ *of smallest degree.*

Solution: The characteristic polynomial is

$$p(x) = (x-3)(x-2)^2 = x^3 - 7x^2 + 16x - 12$$

The minimal polynomial of $\mathbf{A}$ is either

$$(x-3)(x-2) \text{ or } (x-3)(x-2)^2$$

Now

$$(\mathbf{A} - 3\mathbf{I})(\mathbf{A} - 2\mathbf{I}) = \begin{pmatrix} 1 & -2 & 1 \\ 2 & -3 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} 2 & -2 & 1 \\ 2 & -2 & 1 \\ 2 & -2 & 1 \end{pmatrix} = \mathbf{0}$$

Thus the minimal polynomial of $\mathbf{A}$ is $m(x) = (x-3)(x-2) = x^2 - 5x + 6$. Now

$$m(\mathbf{A}) = \mathbf{A}^2 - 5\mathbf{A} + 6\mathbf{I} = \mathbf{0}$$

Hence

$$\begin{aligned} \mathbf{A}^2 &= 5\mathbf{A} - 6\mathbf{I} \\ \mathbf{A}^3 &= 5\mathbf{A}^2 - 6\mathbf{A} \\ &= 5(5\mathbf{A} - 6\mathbf{I}) - 6\mathbf{A} \\ &= 19\mathbf{A} - 30\mathbf{I} \\ \mathbf{A}^4 &= 19\mathbf{A}^2 - 30\mathbf{A} \\ &= 19(5\mathbf{A} - 6\mathbf{I}) - 30\mathbf{A} \\ &= 65\mathbf{A} - 114\mathbf{I} \end{aligned}$$

To find $\mathbf{A}^{-1}$, we have

$$\begin{aligned} \mathbf{A}^2 - 5\mathbf{A} + 6\mathbf{I} &= \mathbf{0} \\ \mathbf{A} - 5\mathbf{I} + 6\mathbf{A}^{-1} &= \mathbf{0} \\ \mathbf{A}^{-1} &= -\frac{1}{6}\mathbf{A} + \frac{5}{6}\mathbf{I} \end{aligned}$$

□

Example 5.4.7. *Let*

$$\mathbf{A} = \begin{pmatrix} 4 & 0 & 4 \\ 0 & 2 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

Find the minimal polynomial of $\mathbf{A}$*. Then express* $\mathbf{A}^4$ *and* $\mathbf{A}^{-1}$ *as a polynomial in* $\mathbf{A}$ *of smallest degree.*

Solution: The characteristic polynomial is

$$p(x) = (x-2)^3$$

The minimal polynomial of $\mathbf{A}$ is either

$$x-2 \quad \text{or} \quad (x-2)^2 \quad \text{or} \quad (x-2)^3$$

Now

$$(\mathbf{A}-2\mathbf{I})^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbf{0}$$

Thus the minimal polynomial of $\mathbf{A}$ is $m(x) = (x-2)^3 = x^3 - 6x^2 + 12x - 8$. Now

$$m(\mathbf{A}) = \mathbf{A}^3 - 6\mathbf{A}^2 + 12\mathbf{A} - 8\mathbf{I} = \mathbf{0}$$

Hence

$$\begin{aligned} \mathbf{A}^3 &= 6\mathbf{A}^2 - 12\mathbf{A} + 8\mathbf{I} \\ \mathbf{A}^4 &= 6\mathbf{A}^3 - 12\mathbf{A}^2 + 8\mathbf{A} \\ &= 6(6\mathbf{A}^2 - 12\mathbf{A} + 8\mathbf{I}) - 12\mathbf{A}^2 + 8\mathbf{A} \\ &= 24\mathbf{A}^2 - 64\mathbf{A} + 48\mathbf{I} \end{aligned}$$

To find $\mathbf{A}^{-1}$, we have

$$\begin{aligned} \mathbf{A}^3 - 6\mathbf{A}^2 + 12\mathbf{A} - 8\mathbf{I} &= \mathbf{0} \\ \mathbf{A}^2 - 6\mathbf{A} + 12\mathbf{I} - 8\mathbf{A}^{-1} &= \mathbf{0} \\ \mathbf{A}^{-1} &= \frac{1}{8}\mathbf{A}^2 - \frac{3}{4}\mathbf{A} + \frac{3}{2}\mathbf{I} \end{aligned}$$

□

Exercise 5.4

1. Find the minimal polynomial of $\mathbf{A}$ where $\mathbf{A}$ is the matrix given below. Then express $\mathbf{A}^4$ and $\mathbf{A}^{-1}$ as a polynomial in $\mathbf{A}$ of smallest degree.

   (a) $\begin{pmatrix} 5 & -4 \\ 3 & -2 \end{pmatrix}$

   (b) $\begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}$

   (c) $\begin{pmatrix} 2 & 5 \\ -1 & 0 \end{pmatrix}$

   (d) $\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$

   (e) $\begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix}$

   (f) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & 1 & 1 \end{pmatrix}$

   (g) $\begin{pmatrix} 11 & -6 & -2 \\ 20 & -11 & -4 \\ 0 & 0 & 1 \end{pmatrix}$

2. Prove that similar matrices have the same minimal polynomial.

3. Let $\mathbf{A}$ be a square matrix such that $\mathbf{A}^k = \mathbf{I}$ for some positive integer k. Prove that $\mathbf{A}$ is diagonalizable.

4. Prove that if $\mathbf{A}$ is a non-singular matrix such that $\mathbf{A}^2$ is diagonalizable, then $\mathbf{A}$ is diagonalizable.

6 Systems of first order linear equations

6.1 Basic properties of systems of first order linear equations

In this chapter, we study systems of first order linear equations

$$\left\{\begin{array}{ccccccccccc} x_1' & = & p_{11}(t)x_1 & + & p_{12}(t)x_2 & + & \cdots & + & p_{1n}(t)x_n & + & g_1(t) \\ x_2' & = & p_{21}(t)x_1 & + & p_{22}(t)x_2 & + & \cdots & + & p_{2n}(t)x_n & + & g_2(t) \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ x_n' & = & p_{n1}(t)x_1 & + & p_{n2}(t)x_2 & + & \cdots & + & p_{nn}(t)x_n & + & g_n(t) \end{array}\right. .$$

where p_{ij}, $g_i(t)$, $i, j = 1, 2, \cdots, n$, are continuous functions. We can also write the system into matrix form

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{g}(t), \quad t \in I,$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{P}(t) = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1n}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}(t) & p_{n2}(t) & \cdots & p_{nn}(t) \end{pmatrix}, \quad \mathbf{g}(t) = \begin{pmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{pmatrix}.$$

An n-th order linear equaition can be transformed to a system of n first order linear equations. Here is an example for second order equation.

Example 6.1.1. *We can use the substitution $x_1(t) = y(t)$ and $x_2(t) = y'(t)$ to transform the second order differential equation*

$$y'' + p(t)y' + q(t)y = g(t),$$

to a system of linear equations

$$\left\{\begin{array}{l} x_1' = x_2 \\ x_2' = -q(t)x_1 - p(t)x_2 + g(t) \end{array}\right. .$$

A fundamental theorem for system of first order linear equations says that a solution always exists and is unique for any given initial condition.

Theorem 6.1.2 (Existence and uniqueness theorem)**.** *If all the functions $\{p_{ij}\}$ and $\{g_i\}$ are continuous on an open interval I, then for any $t_0 \in I$ and $\mathbf{x}_0 \in \mathbb{R}^n$, there exists a unique solution to the initial value problem*

$$\left\{\begin{array}{l} \mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{g}(t), \quad t \in I, \\ \mathbf{x}(t_0) = \mathbf{x}_0. \end{array}\right.$$

Definition 6.1.3. *Let $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ be a set of n solutions to the system $\mathbf{x}' = \mathbf{P}(t)\mathbf{x}$ and let*

$$\mathbf{X}(t) = [\ \mathbf{x}^{(1)} \quad \mathbf{x}^{(2)} \quad \cdots \quad \mathbf{x}^{(n)}\ ],$$

be the $n \times n$ matrix valued function with column vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$. Then the **Wronskian** *of the set of n solutions is defined as*

$$W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t) = \det(\mathbf{X}(t)).$$

The solutions $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ are linearly independent at a point $t_0 \in I$ if and only if $W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t_0) \neq 0$. If $W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t_0) \neq 0$ for some $t_0 \in I$, then we say that $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ form a **fundamental set of solutions**.

Theorem 6.1.4 (Abel's theorem for system of differential equation)**.** *Let* $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \cdots, \mathbf{x}^{(n)}(t)$ *be solutions to the system*

$$\mathbf{x}'(t) = \mathbf{P}(t)\mathbf{x}(t), \quad t \in I.$$

and

$$W(t) = W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t)$$

be the Wronskian. Then $W(t)$ *satisfies the first order linear equation*

$$W'(t) = tr(\mathbf{P})(t)W(t)$$

for some constant c *where* $tr(\mathbf{P})(t) = p_{11}(t)+p_{22}(t)+\cdots+p_{nn}(t)$ *is the trace of* $\mathbf{P}(t)$*. Furthermore* $W(t)$ *is either identically zero on* I *or else never zero on* I*.*

Proof. Differentiating the Wronskian $W(t) = W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t)$ with respect to t, we have

$$\begin{aligned} W' &= \frac{d}{dt}\det\left[\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \mathbf{x}^{(n)}\right] \\ &= \sum_{i=1}^{n}\det\left[\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \frac{d}{dt}\mathbf{x}^{(i)}\ \cdots\mathbf{x}^{(n)}\right] \\ &= \sum_{i=1}^{n}\det\left[\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \mathbf{P}\mathbf{x}^{(i)}\ \cdots\mathbf{x}^{(n)}\right] \\ &= \mathrm{tr}(\mathbf{P})\det\left[\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \mathbf{x}^{(n)}\right] \\ &= \mathrm{tr}(\mathbf{P})W \end{aligned}$$

Here we have used the identity that for any vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n \in \mathbb{R}^n$ and any $n \times n$ matrix $\mathbf{A}$, we have

$$\sum_{i=1}^{n}\det\left[\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{A}\mathbf{x}_i\ \cdots\mathbf{x}_n\right] = \mathrm{tr}(\mathbf{A})\det\left[\mathbf{x}_1\ \mathbf{x}_2\ \cdots\mathbf{x}_n\right]$$

By solving the above first order linear equation for $W(t)$, we have

$$W(t) = c\exp\left(\int \mathrm{tr}(\mathbf{P})(t)dt\right)$$

for some constant c. Now $W(t)$ is identically equal to 0 if c is zero and $W(t)$ is never zero when c is non-zero. □

The above theorem implies that if $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ form a fundamental set of solutions, i.e. $W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t_0) \neq 0$ for some $t_0 \in I$, then $W[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t) \neq 0$ for any $t \in I$ and consequently $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}](t)$ are linearly independent for any $t \in I$.

Theorem 6.1.5. *Suppose* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ *form a fundamental set of solutions to the system*

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x}, \quad t \in I.$$

Then each solution $\mathbf{x}$ *to the system can be expressed as a linear combination*

$$\mathbf{x} = c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \cdots + c_n\mathbf{x}^{(n)}$$

for constants $c_1, c_2, \cdots, c_n$ *in exactly one way.*

Proof. Take an arbitrary $t_0 \in I$. Since $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ form a fundamental set of solutions, we have $\mathbf{x}^{(1)}(t_0), \mathbf{x}^{(2)}(t_0), \cdots, \mathbf{x}^{(n)}(t_0)$ are linearly independent. Thus there exists uniqueness real numbers $c_1, c_2, \cdots, c_n$ such that

$$\mathbf{x}(t_0) = c_1\mathbf{x}^{(1)}(t_0) + c_2\mathbf{x}^{(2)}(t_0) + \cdots + c_n\mathbf{x}^{(n)}(t_0)$$

Now the vector valued function

$$\mathbf{x} - (c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \cdots + c_n\mathbf{x}^{(n)})$$

is also a solution to the system and its value at t_0 is the zero vector $\mathbf{0}$. By uniqueness of solution (Theorem 6.1.2), this vector valued functions is identically equal to zero vector. Therefore $\mathbf{x} = c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \cdots + c_n\mathbf{x}^{(n)}$. This expression is unique because $c_1, c_2, \cdots, c_n$ are unique. □

Exercise 6.1

1. Let $\mathbf{P}(t)$ be a continuous matrix function on an interval I and $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ be solutions to the homogeneous system

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x}, \quad t \in I.$$

   Suppose there exists $t_0 \in I$ such that $\mathbf{x}^{(1)}(t_0), \mathbf{x}^{(2)}(t_0), \cdots, \mathbf{x}^{(n)}(t_0)$ are linearly independent. Show that for any $t \in I$, the vectors $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \cdots, \mathbf{x}^{(n)}(t)$ are linearly independent in $\mathbb{R}^n$.

2. Suppose $\mathbf{x}^{(0)}(t) = (x_1(t), x_2(t), \cdots, x_n(t))^T$ is a solution to the homogeneous system

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x}, \quad t \in I.$$

   Suppose $\mathbf{x}^{(0)}(t_0) = \mathbf{0}$ for some $t_0 \in I$. Show that $\mathbf{x}^{(0)}(t) = \mathbf{0}$ for any $t \in I$.

3. Let $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \cdots, \mathbf{x}^{(n)}(t)$ be differentiable vector valued functions and

$$\mathbf{X}(t) = [\ \mathbf{x}^{(1)} \quad \mathbf{x}^{(2)} \quad \cdots \quad \mathbf{x}^{(n)}\ ],$$

   be the matrix valued function with column vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$. Show that

$$\frac{d}{dt}\det(\mathbf{X}(t)) = \sum_{i=1}^{n} \det\left[\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \frac{d}{dt}\mathbf{x}^{(i)}\ \cdots \mathbf{x}^{(n)}\right]$$

4. Prove that for any vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n \in \mathbb{R}^n$ and any $n \times n$ matrix $\mathbf{A}$, we have

$$\sum_{i=1}^{n} \det\left[\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{A}\mathbf{x}_i\ \cdots \mathbf{x}_n\right] = \operatorname{tr}(\mathbf{A})\det\left[\mathbf{x}_1\ \mathbf{x}_2\ \cdots \mathbf{x}_n\right]$$

6.2 Homogeneous linear systems with constant coefficients

From now on we will consider homogeneous linear systems with constant coefficients

$$\mathbf{x}' = \mathbf{A}\mathbf{x},$$

where $\mathbf{A}$ is a constant $n \times n$ matrix. Suppose the system has a solution of the form

$$\mathbf{x} = e^{\lambda t}\xi,$$

where ξ is a non-zero constant vector. Then

$$\mathbf{x}' = \lambda e^{\lambda t}\xi.$$

Put it into the system, we have

$$(\lambda\mathbf{I} - \mathbf{A})\xi = \mathbf{0}.$$

Since $\xi \neq \mathbf{0}$, λ is an eigenvalue of $\mathbf{A}$ and ξ is an eigenvector associated with λ. Conversely if λ is an eigenvalue of $\mathbf{A}$ and ξ is an eigenvector associated with λ, then $\mathbf{x} = e^{\lambda t}\xi$ gives a solution to the system.

Example 6.2.1. *Solve*

$$\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}\mathbf{x}.$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} \lambda - 1 & -1 \\ -4 & \lambda - 1 \end{vmatrix} &= 0 \\ (\lambda - 1)^2 - 4 &= 0 \\ \lambda - 1 &= \pm 2 \\ \lambda &= 3, -1 \end{aligned}$$

we find that the eigenvalues of the coefficient matrix are $\lambda_1 = 3$ and $\lambda_2 = -1$ and the associated eigenvectors are

$$\xi^{(1)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \xi^{(2)} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

respectively. Therefore the general solution is

$$\mathbf{x} = c_1 e^{3t}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 e^{-t}\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

□

Example 6.2.2. *Solve*

$$\mathbf{x}' = \begin{pmatrix} -3 & \sqrt{2} \\ \sqrt{2} & -2 \end{pmatrix}\mathbf{x}.$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} \lambda + 3 & -\sqrt{2} \\ -\sqrt{2} & \lambda + 2 \end{vmatrix} &= 0 \\ (\lambda + 3)(\lambda + 2) - 2 &= 0 \\ \lambda^2 + 5\lambda + 4 &= \pm 2 \\ \lambda &= -4, -1 \end{aligned}$$

we find that the eigenvalues of the coefficient matrix are $\lambda_1 = -4$ and $\lambda_2 = -1$ and the associated eigenvectors are

$$\xi^{(1)} = \begin{pmatrix} -\sqrt{2} \\ 1 \end{pmatrix}, \quad \xi^{(2)} = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$$

respectively. Therefore the general solution is

$$\mathbf{x} = c_1 e^{-4t}\begin{pmatrix} -\sqrt{2} \\ 1 \end{pmatrix} + c_2 e^{-t}\begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}.$$

☐

When the characteristic equation has repeated root, the above method can still be used if there are n linearly independent eigenvectors, in other words when the coefficient matrix $\mathbf{A}$ is diagonalizable.

Example 6.2.3. *Solve*

$$\mathbf{x}' = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \mathbf{x}.$$

Solution: Solving the characteristic equation, we find that the eigenvalues of the coefficient matrix are $\lambda_1 = 2$ and $\lambda_2 = \lambda_3 = -1$.
For $\lambda_1 = 2$, the associated eigenvector is

$$\xi^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

For the repeated root $\lambda_2 = \lambda_3 = -1$, there are two linearly independent eigenvectors

$$\xi^{(2)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \xi^{(3)} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Therefore the general solution is

$$\mathbf{x} = c_1 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + c_3 e^{-t} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

☐

If $\lambda = \alpha + \beta i, \ \alpha - \beta i, \ \beta > 0$ are complex eigenvalues of $\mathbf{A}$ and $\mathbf{a} + \mathbf{b}i, \ \mathbf{a} - \mathbf{b}i$ are the associated eigenvectors respectively, then the real part and imaginary part of

$$\begin{aligned} e^{(\alpha+\beta i)t}(\mathbf{a} + \mathbf{b}i) &= e^{\alpha t}(\cos \beta t + i \sin \beta t)(\mathbf{a} + \mathbf{b}i) \\ &= e^{\alpha t}(\mathbf{a} \cos \beta t - \mathbf{b} \sin \beta t) + e^{\alpha t}(\mathbf{b} \cos \beta t + \mathbf{a} \sin \beta t)i \end{aligned}$$

give two linearly independent solutions to the system. We have

Theorem 6.2.4. *Suppose* $\lambda = \alpha + \beta i, \ \alpha - \beta i, \ \beta > 0$ *are complex eigenvalues of* $\mathbf{A}$ *and* $\mathbf{a} + \mathbf{b}i, \ \mathbf{a} - \mathbf{b}i$ *are the associated eigenvectors respectively, then*

$$\begin{cases} \mathbf{x}^{(1)} = e^{\alpha t}(\mathbf{a} \cos \beta t - \mathbf{b} \sin \beta t), \\ \mathbf{x}^{(2)} = e^{\alpha t}(\mathbf{b} \cos \beta t + \mathbf{a} \sin \beta t), \end{cases}$$

are two linear independent solutions to $\mathbf{x}' = \mathbf{A}\mathbf{x}$.

Example 6.2.5. *Solve*

$$\mathbf{x}' = \begin{pmatrix} -3 & -2 \\ 4 & 1 \end{pmatrix} \mathbf{x}$$

Solution: Solving the characteristic equation

$$\begin{aligned}
\begin{vmatrix} \lambda+3 & 2 \\ 4 & \lambda-1 \end{vmatrix} &= 0 \\
(\lambda+3)(\lambda-1)+8 &= 0 \\
\lambda^2+2\lambda+5 &= 0 \\
(\lambda+1)^2+4 &= 0 \\
\lambda &= -1 \pm 2i
\end{aligned}$$

For $\lambda_1 = -1+2i$,

$$\mathbf{A} - \lambda_1 \mathbf{I} = \begin{pmatrix} -2+2i & -2 \\ 4 & 2-2i \end{pmatrix}$$

An associated eigenvector is

$$\xi_1 = \begin{pmatrix} -1 \\ 1+i \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} i$$

Moreover an eigenvector associated with $\lambda_2 = -1-2i$ is

$$\xi_1 = \begin{pmatrix} -1 \\ 1-i \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} i$$

Therefore

$$\begin{cases} \mathbf{x}^{(1)} = e^{-t}\left[\begin{pmatrix} -1 \\ 1 \end{pmatrix} \cos 2t - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \sin 2t\right] = e^{-t}\begin{pmatrix} -\cos 2t \\ \cos 2t - \sin 2t \end{pmatrix} \\ \mathbf{x}^{(2)} = e^{-t}\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cos 2t + \begin{pmatrix} -1 \\ 1 \end{pmatrix} \sin 2t\right] = e^{-t}\begin{pmatrix} -\sin 2t \\ \cos 2t + \sin 2t \end{pmatrix} \end{cases}$$

are two linearly independent solutions and the general solution is

$$\begin{aligned}
\mathbf{x} &= c_1 e^{-t}\begin{pmatrix} -\cos 2t \\ \cos 2t - \sin 2t \end{pmatrix} + c_2 e^{-t}\begin{pmatrix} -\sin 2t \\ \cos 2t + \sin 2t \end{pmatrix} \\
&= e^{-t}\begin{pmatrix} -c_1 \cos 2t - c_2 \sin 2t \\ (c_1+c_2)\cos 2t + (c_2-c_1)\sin 2t \end{pmatrix}.
\end{aligned}$$

□

Example 6.2.6. *Two masses m_1 and m_2 are attached to each other and to outside walls by three springs with spring constants k_1, k_2 and k_3 in the straight-line horizontal fashion. Suppose that $m_1 = 2$, $m_2 = 9/4$, $k_1 = 1$, $k_2 = 3$ and $k_3 = 15/4$ Find the displacement of the masses x_1 and x_3 after time t with the initial conditions $x_1(0) = 6$, $x_1'(0) = -6$, $x_2(0) = 4$ and $x_2'(0) = 8$.*

Solution: The equation of motion of the system is

$$\begin{cases} m_1 x_1'' = -(k_1+k_2)x_1 + k_2 x_2, \\ m_2 x_2'' = k_2 x_1 - (k_1+k_3)x_2. \end{cases}$$

Let $y_1 = x_1$, $y_2 = x_2$, $y_3 = x_1'$ and $y_4 = x_2'$. Then the equation is transformed to

$$\mathbf{y}' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 3/2 & 0 & 0 \\ 4/3 & -3 & 0 & 0 \end{pmatrix} \mathbf{y} = \mathbf{A}\mathbf{y}.$$

The characteristic equation is

$$\begin{aligned}
\begin{vmatrix} -\lambda & 0 & 1 & 0 \\ 0 & -\lambda & 0 & 1 \\ -2 & 3/2 & -\lambda & 0 \\ 4/3 & -3 & 0 & -\lambda \end{vmatrix} &= 0 \\
\begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\lambda^2-2 & 3/2 & -\lambda & 0 \\ 4/3 & -\lambda^2-3 & 0 & -\lambda \end{vmatrix} &= 0 \\
(\lambda^2+2)(\lambda^2+3)-2 &= 0 \\
\lambda^4+5\lambda^2+4 &= 0 \\
(\lambda^2+1)(\lambda^2+4) &= 0
\end{aligned}$$

The four eigenvalues are $\lambda_1 = i$, $\lambda_2 = -i$, $\lambda_3 = 2i$ and $\lambda_4 = -2i$. The associated eigenvectors are

$$\xi^{(1)} = \begin{pmatrix} 3 \\ 2 \\ 3i \\ 2i \end{pmatrix}, \quad \xi^{(2)} = \begin{pmatrix} 3 \\ 2 \\ -3i \\ -2i \end{pmatrix}, \quad \xi^{(3)} = \begin{pmatrix} 3 \\ -4 \\ 6i \\ -8i \end{pmatrix}, \quad \xi^{(4)} = \begin{pmatrix} 3 \\ -4 \\ -6i \\ 8i \end{pmatrix}.$$

From real and imaginary parts of

$$\begin{aligned}
e^{\lambda_1 t}\xi^{(1)} &= \begin{pmatrix} 3 \\ 2 \\ 3i \\ 2i \end{pmatrix}(\cos t + i\sin t) \\
&= \begin{pmatrix} 3\cos t \\ 2\cos t \\ -3\sin t \\ -2\sin t \end{pmatrix} + \begin{pmatrix} 3\sin t \\ 2\sin t \\ 3\cos t \\ 2\cos t \end{pmatrix} i
\end{aligned}$$

and

$$\begin{aligned}
e^{\lambda_3 t}\xi^{(3)} &= \begin{pmatrix} 3 \\ -4 \\ 6i \\ -8i \end{pmatrix}(\cos 2t + i\sin 2t) \\
&= \begin{pmatrix} 3\cos 2t \\ -4\cos 2t \\ -6\sin 2t \\ 8\sin 2t \end{pmatrix} + \begin{pmatrix} 3\sin 2t \\ -4\sin 2t \\ 6\cos 2t \\ -8\cos 2t \end{pmatrix} i,
\end{aligned}$$

the general solution to the system is

$$\mathbf{y} = c_1\begin{pmatrix} 3\cos t \\ 2\cos t \\ -3\sin t \\ -2\sin t \end{pmatrix} + c_2\begin{pmatrix} 3\sin t \\ 2\sin t \\ 3\cos t \\ 2\cos t \end{pmatrix} + c_3\begin{pmatrix} 3\cos 2t \\ -4\cos 2t \\ -6\sin 2t \\ 8\sin 2t \end{pmatrix} + c_4\begin{pmatrix} 3\sin 2t \\ -4\sin 2t \\ 6\cos 2t \\ -8\cos 2t \end{pmatrix}.$$

From the initial conditions, we have

$$\begin{aligned}\begin{pmatrix} 3 & 0 & 3 & 0 \\ 2 & 0 & -4 & 0 \\ 0 & 3 & 0 & 6 \\ 0 & 2 & 0 & -8 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} &= \begin{pmatrix} 6 \\ 4 \\ -6 \\ 8 \end{pmatrix} \\ \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} &= \begin{pmatrix} 2 \\ 0 \\ 0 \\ -1 \end{pmatrix}.\end{aligned}$$

Therefore

$$\begin{cases} x_1 &= 6\cos t - 3\sin 2t \\ x_2 &= 4\cos t + 4\sin 2t. \end{cases}$$

□

Exercise 6.2

1. Find the general solutions to the following systems of differential equations.

   (a) $\begin{cases} x_1' = x_1 + 2x_2 \\ x_2' = 2x_1 + x_2 \end{cases}$

   (b) $\begin{cases} x_1' = 2x_1 + 3x_2 \\ x_2' = 2x_1 + x_2 \end{cases}$

   (c) $\begin{cases} x_1' = x_1 - 5x_2 \\ x_2' = x_1 - x_2 \end{cases}$

   (d) $\begin{cases} x_1' = 5x_1 - 9x_2 \\ x_2' = 2x_1 - x_2 \end{cases}$

   (e) $\begin{cases} x_1' = 4x_1 + x_2 + 4x_3 \\ x_2' = x_1 + 7x_2 + x_3 \\ x_3' = 4x_1 + x_2 + 4x_3 \end{cases}$

   (f) $\begin{cases} x_1' = 4x_1 + x_2 + x_3 \\ x_2' = x_1 + 4x_2 + x_3 \\ x_3' = x_1 + x_2 + 4x_3 \end{cases}$

   (g) $\begin{cases} x_1' = 2x_1 + x_2 - x_3 \\ x_2' = -4x_1 - 3x_2 - x_3 \\ x_3' = 4x_1 + 4x_2 + 2x_3 \end{cases}$

   (h) $\begin{cases} x_1' = 2x_1 + 2x_2 + x_3 \\ x_2' = x_1 + 3x_2 + x_3 \\ x_3' = x_1 + 2x_2 + 2x_3 \end{cases}$

2. Solve the following initial value problem.

   (a) $\begin{cases} x_1' = 3x_1 + 4x_2 \\ x_2' = 3x_1 + 2x_2 \\ x_1(0) = x_2(0) = 1 \end{cases}$

   (b) $\begin{cases} x_1' = 9x_1 + 5x_2 \\ x_2' = -6x_1 - 2x_2 \\ x_1(0) = 1,\ x_2(0) = 0 \end{cases}$

   (c) $\begin{cases} x_1' = x_1 - 2x_2 \\ x_2' = 2x_1 + x_2 \\ x_1(0) = 0,\ x_2(0) = 4 \end{cases}$

   (d) $\begin{cases} x_1' = 3x_1 + x_3 \\ x_2' = 9x_1 - x_2 + 2x_3 \\ x_3' = -9x_1 + 4x_2 - x_3 \\ x_1(0) = x_2(0) = 0,\ x_3(0) = 17 \end{cases}$

3. Solve $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for the given matrix $\mathbf{A}$.

(a) $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$

(b) $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 5 & -1 \end{pmatrix}$

(c) $\mathbf{A} = \begin{pmatrix} 2 & -3 & 3 \\ 4 & -5 & 3 \\ 4 & -4 & 2 \end{pmatrix}$

(d) $\mathbf{A} = \begin{pmatrix} 4 & -1 & -1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$

6.3 Repeated eigenvalues

Consider the system

$$\mathbf{x}' = \mathbf{A}\mathbf{x}, \quad \text{where } \mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}.$$

The characteristic equation of $\mathbf{A}$ is

$$\begin{aligned} \begin{vmatrix} \lambda - 1 & 1 \\ -1 & \lambda - 3 \end{vmatrix} &= 0 \\ \lambda^2 - 4\lambda + 4 &= 0 \\ (\lambda - 2)^2 &= 0 \end{aligned}$$

It has a root $\lambda = 2$ of multiplicity $m_a = 2$. However, the set of all eigenvectors associated with $\lambda = 2$ is spanned by one vector

$$\xi = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

In other words, the geometric multiplicity $\lambda = 2$ is $m_g = 1$. So the geometric multiplicity is smaller than the algebraic multiplicity. We know that

$$\mathbf{x}^{(1)} = e^{\lambda t}\xi = \begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix}$$

is a solution to the system. However we do not have sufficient number of linearly independent eigenvectors to write down two linearly independent solutions to the system. How do we find another solution to form a fundamental set of solutions?

Based on the procedure used for higher order linear equations, it may be natural to attempt to find a second solution of the form

$$\mathbf{x} = te^{2t}\xi.$$

Substituting this into the equation reads

$$\begin{aligned} \frac{d}{dt}te^{2t}\xi &= \mathbf{A}te^{2t}\xi \\ 2te^{2t}\xi + e^{2t}\xi &= te^{2t}\mathbf{A}\xi \\ e^{2t}\xi &= te^{2t}(\mathbf{A} - 2\mathbf{I})\xi \\ e^{2t}\xi &= \mathbf{0} \end{aligned}$$

which has no non-zero solution for eigenvector ξ. To overcome this problem, we try another substitution

$$\mathbf{x} = te^{2t}\xi + e^{2t}\eta,$$

where η is a vector to be determined. Then the equation $\mathbf{x}' = \mathbf{A}\mathbf{x}$ reads

$$\begin{aligned} 2te^{2t}\xi + e^{2t}(\xi + 2\eta) &= \mathbf{A}(te^{2t}\xi + e^{2t}\eta) \\ e^{2t}\xi &= e^{2t}(\mathbf{A} - 2\mathbf{I})\eta \\ \xi &= (\mathbf{A} - 2\mathbf{I})\eta \end{aligned}$$

Therefore if we take η such that $\xi = (\mathbf{A} - 2\mathbf{I})\eta$ is an eigenvector, then $\mathbf{x} = te^{2t}\xi + e^{2t}\eta$ is another solution to the system. Note that η satisfies

$$\begin{cases} (\mathbf{A} - \lambda\mathbf{I})\eta \neq \mathbf{0} \\ (\mathbf{A} - \lambda\mathbf{I})^2\eta = \mathbf{0} \end{cases}$$

A vector satisfying these two equations is called a **generalized eigenvector** of **rank** 2 associated with eigenvalue λ. Back to our example, if we take

$$\eta = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Then

$$\begin{cases} (\mathbf{A} - 2\mathbf{I})\eta = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \neq \mathbf{0} \\ (\mathbf{A} - 2\mathbf{I})^2\eta = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \mathbf{0} \end{cases}$$

So η is a generalized eigenvector of rank 2. Then

$$\mathbf{x}^{(2)} = te^{2t}\xi + e^{2t}\eta = te^{2t}(\mathbf{A} - 2\mathbf{I})\eta + e^{2t}\eta = te^{2t}\begin{pmatrix} -1 \\ 1 \end{pmatrix} + e^{2t}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

is a solution to the system and the general solution is

$$\begin{aligned} \mathbf{x} &= c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} \\ &= c_1e^{2t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2\left(te^{2t}\begin{pmatrix} -1 \\ 1 \end{pmatrix} + e^{2t}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \\ &= e^{2t}\begin{pmatrix} c_1 + c_2 - c_2t \\ -c_1 + c_2t \end{pmatrix}. \end{aligned}$$

In the above example, we see how generalized eigenvectors can to used to write down more solutions when the coefficient matrix of the system is not diagonalizable.

Definition 6.3.1 (Generalized eigenvector)**.** *Let* $\mathbf{A}$ *be a square matrix,* λ *be an eigenvalue of* $\mathbf{A}$ *and* k *be a positive integer. A non-zero vector* η *is called a* **generalized eigenvector of rank** k *associated with eigenvalue* λ *if*

$$\begin{cases} (\mathbf{A} - \lambda\mathbf{I})^{k-1}\eta \neq \mathbf{0}, \\ (\mathbf{A} - \lambda\mathbf{I})^k\eta = \mathbf{0}. \end{cases}$$

Note that a vector is a generalized eigenvector of rank 1 if and only if it is an ordinary eigenvector.

Theorem 6.3.2. *Let* $\mathbf{A}$ *be a square matrix and* η *be a generalized eigenvector of rank* k *associated with eigenvalue* λ*. Let*

$$\begin{cases} \eta_0 &= \eta, \\ \eta_1 &= (\mathbf{A} - \lambda\mathbf{I})\eta, \\ \eta_2 &= (\mathbf{A} - \lambda\mathbf{I})^2\eta, \\ &\vdots \\ \eta_{k-1} &= (\mathbf{A} - \lambda\mathbf{I})^{k-1}\eta, \\ \eta_k &= (\mathbf{A} - \lambda\mathbf{I})^k\eta = \mathbf{0}. \end{cases}$$

Then

1. *For $0 \leq i \leq k-1$, we have η_i is a generalized eigenvector of rank $k-i$ associated with eigenvalue λ.*
2. *The vectors $\eta, \eta_1, \eta_2, \cdots, \eta_{k-1}$ are linearly independent.*

Proof. It is easy to see that $\eta_i = (\mathbf{A} - \lambda\mathbf{I})^i\eta$ satisfies

$$\begin{cases} (\mathbf{A} - \lambda\mathbf{I})^{k-i-1}\eta_i = (\mathbf{A} - \lambda\mathbf{I})^{k-1}\eta = \eta_{k-1} \neq \mathbf{0}, \\ (\mathbf{A} - \lambda\mathbf{I})^{k-i}\eta_i = (\mathbf{A} - \lambda\mathbf{I})^{k}\eta = \eta_k = \mathbf{0} \end{cases}$$

and the first statement follows. We prove the second statement by induction on k. The theorem is obvious when $k = 1$ since η is non-zero. Assume that the theorem is true for rank of η small than k. Suppose η is a generalized eigenvector of rank k associated with eigenvalue λ and $c_0, c_1, \cdots, c_{k-1}$ are scalars such that

$$c_0\eta + c_1\eta_1 + \cdots + c_{k-2}\eta_{k-2} + c_{k-1}\eta_{k-1} = \mathbf{0}.$$

Multiplying both sides from the left by $\mathbf{A} - \lambda\mathbf{I}$, we have

$$c_0\eta_1 + c_1\eta_2 + \cdots + c_{k-2}\eta_{k-1} = \mathbf{0}.$$

Here we used $\eta_k = (\mathbf{A} - \lambda\mathbf{I})\eta_{k-1} = \mathbf{0}$. Now η_1 is a generalized eigenvector of rank $k-1$ by the first statement. Thus by induction hypothesis, we have $\eta_1, \eta_2, \cdots, \eta_{k-1}$ are linearly independent and hence

$$c_0 = c_1 = \cdots = c_{k-2} = 0.$$

Combining the first equality gives $c_{k-1}\eta_{k-1} = \mathbf{0}$ which implies $c_{k-1} = 0$ since η_{k-1} is non-zero. We conclude that $\eta, \eta_1, \eta_2, \cdots, \eta_{k-1}$ are linearly independent. □

A generalized eigenvector of rank k can be used to write down k linearly independent eigenvectors for the system.

Theorem 6.3.3. *Suppose λ is an eigenvalue of a $n \times n$ matrix $\mathbf{A}$ and η is a generalized eigenvector of rank k associated with λ. For $i = 0, 1, 2, \cdots, k-1$, define*

$$\eta_i = (\mathbf{A} - \lambda\mathbf{I})^i\eta$$

Then

$$\begin{cases} \mathbf{x}^{(1)} &= e^{\lambda t}\eta_{k-1} \\ \mathbf{x}^{(2)} &= e^{\lambda t}(\eta_{k-2} + t\eta_{k-1}) \\ \mathbf{x}^{(3)} &= e^{\lambda t}(\eta_{k-3} + t\eta_{k-2} + \frac{t^2}{2}\eta_{k-1}) \\ &\vdots \\ \mathbf{x}^{(k)} &= e^{\lambda t}(\eta + t\eta_1 + \frac{t^2}{2}\eta_2 + \cdots + \frac{t^{k-2}}{(k-2)!}\eta_{k-2} + \frac{t^{k-1}}{(k-1)!}\eta_{k-1}) \end{cases}$$

are linearly independent solutions to the system

$$\mathbf{x}' = \mathbf{A}\mathbf{x}.$$

Proof. It is left for the reader to check that the solutions are linearly independent. It suffices to prove that $\mathbf{x}^{(k)}$ is a solution to the system. Observe that for any $0 \leq i \leq k-1$,

$$\begin{aligned} \mathbf{A}\eta_i &= \lambda\eta_i + (\mathbf{A} - \lambda\mathbf{I})\eta_i \\ &= \lambda\eta_i + \eta_{i+1}. \end{aligned}$$

Thus

$$\begin{aligned}
&\frac{d\mathbf{x}^{(k)}}{dt}\\
= \;& \frac{d}{dt}\left(e^{\lambda t}(\eta + t\eta_1 + \frac{t^2}{2}\eta_2 + \cdots + \frac{t^{k-1}}{(k-1)!}\eta_{k-1})\right)\\
= \;& e^{\lambda t}\left(\lambda\eta + (1+\lambda t)\eta_1 + (t + \frac{\lambda t^2}{2})\eta_2 + \cdots + (\frac{t^{k-2}}{(k-2)!} + \frac{\lambda t^{k-1}}{(k-1)!})\eta_{k-1}\right)\\
= \;& e^{\lambda t}\left((\lambda\eta + \eta_1) + t(\lambda\eta_1 + \eta_2) + \frac{t^2}{2}(\lambda\eta_2 + \eta_3) + \cdots + \frac{t^{k-2}}{(k-2)!}(\lambda\eta_{k-2} + \eta_{k-1}) + \frac{t^{k-1}}{(k-1)!}\lambda\eta_{k-1}\right)\\
= \;& e^{\lambda t}\left(\mathbf{A}\eta + t\mathbf{A}\eta_1 + \frac{t^2}{2}\mathbf{A}\eta_2 + \cdots + \frac{t^{k-2}}{(k-2)!}\mathbf{A}\eta_{k-2} + \frac{t^{k-1}}{(k-1)!}\mathbf{A}\eta_{k-1}\right)\\
= \;& \mathbf{A}e^{\lambda t}\left(\eta + t\eta_1 + \frac{t^2}{2}\eta_2 + \cdots + \frac{t^{k-2}}{(k-2)!}\eta_{k-2} + \frac{t^{k-1}}{(k-1)!}\eta_{k-1}\right)\\
= \;& \mathbf{A}\mathbf{x}^{(k)}
\end{aligned}$$

□

Example 6.3.4. *Solve*

$$\mathbf{x}' = \begin{pmatrix} 1 & -3 \\ 3 & 7 \end{pmatrix}\mathbf{x}.$$

Solution: The characteristic equation of the coefficient matrix **A** is

$$\begin{aligned}
\begin{vmatrix} \lambda - 1 & 3 \\ -3 & \lambda - 7 \end{vmatrix} &= 0\\
(\lambda - 4)^2 &= 0.
\end{aligned}$$

We find that $\lambda = 4$ is double root and the eigenspace associated with $\lambda = 4$ is of dimension 1 and is spanned by $(1, -1)^T$. Thus

$$\mathbf{x}^{(1)} = e^{4t}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

is a solution. To find another solution which is not a multiple of $\mathbf{x}^{(1)}$, we need to find a generalized eigenvector of rank 2. First we calculate

$$\mathbf{A} - 4\mathbf{I} = \begin{pmatrix} -3 & -3 \\ 3 & 3 \end{pmatrix}.$$

Now we if take

$$\eta = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

then η satisfies

$$\begin{cases} \eta_1 = (\mathbf{A} - 4\mathbf{I})\eta = \begin{pmatrix} -3 \\ 3 \end{pmatrix} \neq \mathbf{0},\\ \eta_2 = (\mathbf{A} - 4\mathbf{I})^2\eta = \mathbf{0}. \end{cases}$$

Thus η is a generalized eigenvector of rank 2. Hence

$$\begin{aligned}
\mathbf{x}^{(2)} &= e^{\lambda t}(\eta + t\eta_1)\\
&= e^{4t}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -3 \\ 3 \end{pmatrix}\right)\\
&= e^{4t}\begin{pmatrix} 1 - 3t \\ 3t \end{pmatrix}
\end{aligned}$$

is another solution to the system. Therefore the general solution is

$$\begin{aligned}\mathbf{x} &= c_1e^{4t}\begin{pmatrix}1\\-1\end{pmatrix}+c_2e^{4t}\begin{pmatrix}1-3t\\3t\end{pmatrix}\\ &= e^{4t}\begin{pmatrix}c_1+c_2-3c_2t\\-c_1+3c_2t\end{pmatrix}\end{aligned}$$

□

Example 6.3.5. *Solve*

$$\mathbf{x}'=\begin{pmatrix}7&1\\-4&3\end{pmatrix}\mathbf{x}.$$

Solution: The characteristic equation of the coefficient matrix **A** is

$$\begin{aligned}\begin{vmatrix}\lambda-7&-1\\4&\lambda-3\end{vmatrix} &= 0\\ (\lambda-5)^2 &= 0.\end{aligned}$$

We find that $\lambda=5$ is double root and the eigenspace associated with $\lambda=5$ is of dimension 1 and is spanned by $(1,-2)^T$. Thus

$$\mathbf{x}^{(1)}=e^{5t}\begin{pmatrix}1\\-2\end{pmatrix}$$

is a solution. To find the second solution, we calculate

$$\mathbf{A}-5\mathbf{I}=\begin{pmatrix}2&1\\-4&-2\end{pmatrix}.$$

Now we if take

$$\eta=\begin{pmatrix}1\\0\end{pmatrix},$$

then η satisfies

$$\begin{cases}\eta_1=(\mathbf{A}-5\mathbf{I})\eta=\begin{pmatrix}2\\-4\end{pmatrix}\neq\mathbf{0},\\ \eta_2=(\mathbf{A}-5\mathbf{I})^2\eta=(\mathbf{A}-5\mathbf{I})\begin{pmatrix}2\\-4\end{pmatrix}=\mathbf{0}.\end{cases}$$

Thus η is a generalized eigenvector of rank 2. Hence

$$\begin{aligned}\mathbf{x}^{(2)} &= e^{\lambda t}(\eta+t\eta_1)\\ &= e^{5t}\left(\begin{pmatrix}1\\0\end{pmatrix}+t\begin{pmatrix}2\\-4\end{pmatrix}\right)\\ &= e^{5t}\begin{pmatrix}1+2t\\-4t\end{pmatrix}\end{aligned}$$

is another solution to the system. Therefore the general solution is

$$\begin{aligned}\mathbf{x} &= c_1e^{5t}\begin{pmatrix}1\\-2\end{pmatrix}+c_2e^{5t}\begin{pmatrix}1+2t\\-4t\end{pmatrix}\\ &= e^{5t}\begin{pmatrix}c_1+c_2+2c_2t\\-2c_1-4c_2t\end{pmatrix}\end{aligned}$$

□

Example 6.3.6. *Solve*

$$\mathbf{x}' = \begin{pmatrix} 0 & 1 & 2 \\ -5 & -3 & -7 \\ 1 & 0 & 0 \end{pmatrix} \mathbf{x}.$$

Solution: The characteristic equation of the coefficient matrix $\mathbf{A}$ is

$$\begin{aligned} \begin{vmatrix} \lambda & -1 & -2 \\ 5 & \lambda+3 & 7 \\ -1 & 0 & \lambda \end{vmatrix} &= 0 \\ (\lambda+1)^3 &= 0. \end{aligned}$$

Thus $\mathbf{A}$ has an eigenvalue $\lambda = -1$ of multiplicity 3. By considering

$$\mathbf{A} + \mathbf{I} = \begin{pmatrix} 1 & 1 & 2 \\ -5 & -2 & -7 \\ 1 & 0 & 1 \end{pmatrix}$$

we see that the associated eigenspace is of dimension 1 and is spanned by $(1, 1, -1)^T$. We need to find a generalized eigenvector of rank 3, that is, a vector η such that

$$\begin{cases} (\mathbf{A}+\mathbf{I})^2\eta \neq \mathbf{0} \\ (\mathbf{A}+\mathbf{I})^3\eta = \mathbf{0} \end{cases}$$

Note that by Cayley-Hamilton Theorem, we have $(\mathbf{A}+\mathbf{I})^3 = \mathbf{0}$. Thus the condition $(\mathbf{A}+\mathbf{I})^3\eta = \mathbf{0}$ is automatic. We need to find η which satisfies the first condition. Now we take $\eta = (1, 0, 0)^T$, then

$$\begin{cases} \eta_1 = (\mathbf{A}+\mathbf{I})\eta = \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} \neq \mathbf{0}, \\ \eta_2 = (\mathbf{A}+\mathbf{I})^2\eta = \begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix} \neq \mathbf{0}. \end{cases}$$

(One may verify that $(\mathbf{A}+\mathbf{I})^3\eta = \mathbf{0}$ though it is automatic.) Therefore ξ is a generalized eigenvector of rank 3 associated with $\lambda = -1$. Hence

$$\begin{cases} \mathbf{x}^{(1)} = e^{\lambda t}\eta_2 = e^{-t}\begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix} \\ \mathbf{x}^{(2)} = e^{\lambda t}(\eta_1 + t\eta_2) = e^{-t}\begin{pmatrix} 1-2t \\ -5-2t \\ 1+2t \end{pmatrix} \\ \mathbf{x}^{(3)} = e^{\lambda t}(\eta + t\eta_1 + \frac{t^2}{2}\eta_2) = e^{-t}\begin{pmatrix} 1+t-t^2 \\ -5t-t^2 \\ t+t^2 \end{pmatrix} \end{cases}$$

form a fundamental set of solutions to the system. □

Example 6.3.7. *Solve*

$$\mathbf{x}' = \begin{pmatrix} 0 & 1 & -2 \\ 8 & -1 & 6 \\ 7 & -3 & 8 \end{pmatrix} \mathbf{x}.$$

Solution: The characteristic equation of the coefficient matrix $\mathbf{A}$

$$\begin{aligned}\begin{vmatrix} \lambda & -1 & 2 \\ -8 & \lambda+1 & -6 \\ -7 & 3 & \lambda-8 \end{vmatrix} &= 0 \\ (\lambda-3)(\lambda-2)^2 &= 0\end{aligned}$$

has a simple root $\lambda_1 = 3$ and a double root $\lambda_2 = \lambda_3 = 2$. The eigenspace associated with the simple root $\lambda_1 = 3$ is of dimension 1 and is spanned by $(-1,1,2)^T$. The eigenspace associated with the double root $\lambda_2 = \lambda_3 = 2$ is of dimension 1 and is spanned by $(0,2,1)^T$. We obtain two linearly independent solutions

$$\begin{cases} \mathbf{x}^{(1)} = e^{3t}\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \\ \mathbf{x}^{(2)} = e^{2t}\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \end{cases}.$$

To find the third solution, we need to find a generalized eigenvector of rank 2 associated with the double root $\lambda_2 = \lambda_3 = 2$. The null space of

$$(\mathbf{A}-2\mathbf{I})^2 = \begin{pmatrix} -2 & 1 & -2 \\ 8 & -3 & 6 \\ 7 & -3 & 6 \end{pmatrix}^2 = \begin{pmatrix} -2 & 1 & -2 \\ 2 & -1 & 2 \\ 4 & -2 & 4 \end{pmatrix}$$

is spanned by $(0,2,1)^T$ and $(1,0,-1)^T$. One may check that $(0,2,1)^T$ is an ordinary eigenvector. Now

$$\begin{cases} (\mathbf{A}-2\mathbf{I})\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \neq \mathbf{0} \\ (\mathbf{A}-2\mathbf{I})^2\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 & 1 & -2 \\ 8 & -3 & 6 \\ 7 & -3 & 6 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \mathbf{0} \end{cases}$$

Thus $(1,0,-1)^T$ is a generalized eigenvector of rank 2 associated with $\lambda = 2$. Let

$$\begin{cases} \eta = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ \eta_1 = (\mathbf{A}-2\mathbf{I})\eta = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \end{cases}.$$

We obtain the third solution

$$\mathbf{x}^{(3)} = e^{\lambda t}(\eta + t\eta_1) = e^{2t}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}\right) = e^{2t}\begin{pmatrix} 1 \\ 2t \\ -1+t \end{pmatrix}$$

Therefore the general solution is

$$\mathbf{x} = c_1e^{3t}\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} + c_2e^{2t}\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} + c_3e^{2t}\begin{pmatrix} 1 \\ 2t \\ -1+t \end{pmatrix}.$$

□

Exercise 6.3

1. Find the general solution to the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for the given matrix $\mathbf{A}$.

   (a) $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix}$

   (b) $\mathbf{A} = \begin{pmatrix} -2 & 1 \\ -1 & -4 \end{pmatrix}$

   (c) $\mathbf{A} = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$

   (d) $\mathbf{A} = \begin{pmatrix} -3 & 0 & -4 \\ -1 & -1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$

   (e) $\mathbf{A} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & -4 \\ 0 & 1 & -3 \end{pmatrix}$

   (f) $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & -2 & -3 \\ 2 & 3 & 4 \end{pmatrix}$

   (g) $\mathbf{A} = \begin{pmatrix} 3 & -1 & -3 \\ 1 & 1 & -3 \\ 0 & 0 & 2 \end{pmatrix}$

   (h) $\mathbf{A} = \begin{pmatrix} 3 & 1 & -1 \\ -1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

6.4 Matrix exponential

Let $\mathbf{A}$ be an $n \times n$ matrix. We are going to define $e^{\mathbf{A}}$. This cannot be interpreted as 'e to the power $\mathbf{A}$' because we do not know how to define raising to the power of a matrix. However the exponential function is defined by

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

and the right hand side is defined when x is a square matrix. This inspires us to make the following definition.

Definition 6.4.1 (Matrix exponential). *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. The* **matrix exponential** *of* $\mathbf{A}$ *is defined as*

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!} = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \cdots$$

The matrix exponential of $\mathbf{A}$ *may also be denoted by* $e^{\mathbf{A}}$.

For the purpose of solving system of differential equations, it is helpful to consider matrix valued function

$$\exp(\mathbf{A}t) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!} = \mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \frac{1}{3!}\mathbf{A}^3 t^3 + \cdots$$

for constant square matrix $\mathbf{A}$ and real variable t. It is not difficult to calculate $\exp(\mathbf{A}t)$ if $\mathbf{A}$ is diagonalizable.

Theorem 6.4.2. *Suppose*

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & & & \mathit{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathit{0} & & & \lambda_n \end{pmatrix}$$

is a diagonal matrix. Then

$$\exp(\mathbf{D}t) = \begin{pmatrix} e^{\lambda_1 t} & & & \mathit{0} \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ \mathit{0} & & & e^{\lambda_n t} \end{pmatrix}.$$

Moreover, for any $n \times n$ matrix $\mathbf{A}$, if there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ is a diagonal matrix. Then

$$\exp(\mathbf{A}t) = \mathbf{P}\exp(\mathbf{D}t)\mathbf{P}^{-1}.$$

Example 6.4.3. *Find* $\exp(\mathbf{A}t)$ *where*

$$\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}.$$

Solution: Diagonalizing $\mathbf{A}$, we have

$$\begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}^{-1} \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}.$$

Therefore

$$\begin{aligned} \exp(\mathbf{A}t) &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} e^{5t} & 0 \\ 0 & e^{-2t} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} e^{5t} & 0 \\ 0 & e^{-2t} \end{pmatrix} \frac{1}{7} \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix} \\ &= \frac{1}{7} \begin{pmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{pmatrix} \end{aligned}$$

□

Theorem 6.4.4. *Suppose there exists positive integer k such that $\mathbf{A}^k = \mathbf{0}$, then*

$$\exp(\mathbf{A}t) = \mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \frac{1}{3!}\mathbf{A}^3 t^3 + \cdots + \frac{1}{(k-1)!}\mathbf{A}^{k-1} t^{k-1}$$

Proof. It follows easily from the fact that $\mathbf{A}^l = \mathbf{0}$ for all $l \geq k$. □

Example 6.4.5. *Find* $\exp(\mathbf{A}t)$ *where*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Solution: First compute

$$\mathbf{A}^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{A}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore

$$\begin{aligned}\exp(\mathbf{A}t) &= \mathbf{I}+\mathbf{A}t+\frac{1}{2}\mathbf{A}^2t^2\\ &= \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}+t\begin{pmatrix}0&1&3\\0&0&2\\0&0&0\end{pmatrix}+\frac{t^2}{2}\begin{pmatrix}0&0&2\\0&0&0\\0&0&0\end{pmatrix}\\ &= \begin{pmatrix}1&t&3t+2t^2\\0&1&2t\\0&0&1\end{pmatrix}.\end{aligned}$$

□

Theorem 6.4.6. *Let* $\mathbf{A}$ *be an* $n\times n$ *matrix. Then*

$$\frac{d}{dt}\exp(\mathbf{A}t)=\mathbf{A}\exp(\mathbf{A}t)$$

Proof.

$$\begin{aligned}\frac{d}{dt}\exp(\mathbf{A}t) &= \frac{d}{dt}\left(\mathbf{I}+\mathbf{A}t+\frac{1}{2!}\mathbf{A}^2t^2+\frac{1}{3!}\mathbf{A}^3t^3+\cdots\right)\\ &= \mathbf{0}+\mathbf{A}+\frac{1}{2!}\mathbf{A}^2(2t)+\frac{1}{3!}\mathbf{A}^3(3t^2)+\cdots\\ &= \mathbf{A}+\mathbf{A}^2t+\frac{1}{2!}\mathbf{A}^3t^2+\cdots\\ &= \mathbf{A}\left(\mathbf{I}+\mathbf{A}t+\frac{1}{2!}\mathbf{A}^2t^2+\cdots\right)\\ &= \mathbf{A}\exp(\mathbf{A}t)\end{aligned}$$

□

The above theorem implies that the column vectors of $\exp(\mathbf{A}t)$ satisfies the system of first order linear differential equations $\mathbf{x}'=\mathbf{A}\mathbf{x}$.

Theorem 6.4.7. *Let* $\mathbf{A}$ *be an* $n\times n$ *matrix and write* $\exp(\mathbf{A}t)=[\mathbf{x}_1(t)\ \mathbf{x}_2(t)\ \cdots\ \mathbf{x}_n(t)]$*, which means* $\mathbf{x}_1(t),\mathbf{x}_2(t),\cdots,\mathbf{x}_n(t)$ *are the column vectors of* $\exp(\mathbf{A}t)$*. Then* $\mathbf{x}_1(t),\mathbf{x}_2(t),\cdots,\mathbf{x}_n(t)$ *are solutions to the system* $\mathbf{x}'=\mathbf{A}\mathbf{x}$*.*

Proof. By Theorem 6.4.6, we have

$$\begin{aligned}\frac{d}{dt}\exp(\mathbf{A}t) &= \mathbf{A}\exp(\mathbf{A}t)\\ \frac{d}{dt}[\mathbf{x}_1(t)\ \mathbf{x}_2(t)\ \cdots\ \mathbf{x}_n(t)] &= \mathbf{A}[\mathbf{x}_1(t)\ \mathbf{x}_2(t)\ \cdots\ \mathbf{x}_n(t)]\\ [\mathbf{x}_1'(t)\ \mathbf{x}_2'(t)\ \cdots\ \mathbf{x}_n'(t)] &= [\mathbf{A}\mathbf{x}_1(t)\ \mathbf{A}\mathbf{x}_2(t)\ \cdots\ \mathbf{A}\mathbf{x}_n(t)]\end{aligned}$$

Therefore $\mathbf{x}_k'=\mathbf{A}\mathbf{x}_k$ for $k=1,2,\cdots,n$. □

Theorem 6.4.8 (Properties of matrix exponential)**.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $n\times n$ *matrices and* a,b *be any scalars. Then the following statements hold.*

1. $\exp(\mathbf{0})=\mathbf{I}$

2. $\exp((a+b)\mathbf{A}t) = \exp(a\mathbf{A}t)\exp(b\mathbf{A}t)$
3. $\exp(-\mathbf{A}t) = (\exp(\mathbf{A}t))^{-1}$
4. *If* $\mathbf{AB} = \mathbf{BA}$*, then* $\exp((\mathbf{A}+\mathbf{B})t) = \exp(\mathbf{A}t)\exp(\mathbf{B}t)$.
5. *For any non-singular matrix* $\mathbf{P}$*, we have* $\exp(\mathbf{P}^{-1}\mathbf{AP}t) = \mathbf{P}^{-1}\exp(\mathbf{A}t)\mathbf{P}$.
6. $\det(\exp(\mathbf{A}t)) = e^{\mathrm{tr}(\mathbf{A}t)}$. *(*$\mathrm{tr}(\mathbf{A}t) = (a_{11}+a_{22}+\cdots+a_{nn})t$ *is the trace of* $\mathbf{A}t$*.)*

Proof. 1. $\exp(\mathbf{0}) = \mathbf{I} + \mathbf{0} + \frac{1}{2!}\mathbf{0}^2 + \frac{1}{3!}\mathbf{0}^3 + \cdots = \mathbf{I}$

2.

$$\begin{aligned}
\exp((a+b)\mathbf{A}t) &= \sum_{k=0}^{\infty} \frac{(a+b)^k\mathbf{A}^kt^k}{k!} \\
&= \sum_{k=0}^{\infty}\left(\sum_{i=0}^{k}\frac{k!a^ib^{k-i}}{i!(k-i)!}\right)\frac{\mathbf{A}^kt^k}{k!} \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{a^ib^{k-i}\mathbf{A}^kt^k}{i!(k-i)!} \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{a^ib^j\mathbf{A}^{i+j}t^{i+j}}{i!j!} \\
&= \sum_{i=0}^{\infty}\frac{a^i\mathbf{A}^it^i}{i!}\sum_{j=0}^{\infty}\frac{b^j\mathbf{A}^jt^j}{j!} \\
&= \exp(a\mathbf{A}t)\exp(b\mathbf{A}t)
\end{aligned}$$

Here we have changed the order of summation of an infinite series in the fourth line and we can do so because the exponential series is absolutely convergent.

3. From the first and second parts, we have $\exp(\mathbf{A}t)\exp(-\mathbf{A}t) = \exp((t-t)\mathbf{A}) = \exp(\mathbf{0}) = \mathbf{I}$. Thus $\exp(-\mathbf{A}t) = (\exp(\mathbf{A}t))^{-1}$.

4.

$$\begin{aligned}
\exp(t(\mathbf{A}+\mathbf{B})) &= \sum_{k=0}^{\infty}\frac{(\mathbf{A}+\mathbf{B})^kt^k}{k!} \\
&= \sum_{k=0}^{\infty}\left(\sum_{i=0}^{k}\frac{k!\mathbf{A}^i\mathbf{B}^{k-i}}{i!(k-i)!}\right)\frac{t^k}{k!} \quad \text{(We used } \mathbf{AB} = \mathbf{BA}\text{)} \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{\mathbf{A}^i\mathbf{B}^{k-i}t^k}{i!(k-i)!} \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\mathbf{A}^i\mathbf{B}^jt^{i+j}}{i!j!} \quad \text{(We changed the order of summation)} \\
&= \sum_{i=0}^{\infty}\frac{\mathbf{A}^it^i}{i!}\sum_{j=0}^{\infty}\frac{\mathbf{B}^jt^j}{j!} \\
&= \exp(\mathbf{A}t)\exp(\mathbf{B}t)
\end{aligned}$$

5.

$$\begin{aligned}\exp(\mathbf{P}^{-1}\mathbf{AP}t) &= \mathbf{I}+\mathbf{P}^{-1}\mathbf{AP}t+\frac{1}{2!}(\mathbf{P}^{-1}\mathbf{AP})^2t^2+\frac{1}{3!}(\mathbf{P}^{-1}\mathbf{AP})^3t^3+\cdots\\ &= \mathbf{P}^{-1}\mathbf{IP}+\mathbf{P}^{-1}(\mathbf{A}t)\mathbf{P}+\mathbf{P}^{-1}\frac{\mathbf{A}^2t^2}{2!}\mathbf{P}+\mathbf{P}^{-1}\frac{\mathbf{A}^3t^3}{3!}\mathbf{P}+\cdots\\ &= \mathbf{P}^{-1}\left(\mathbf{I}+\mathbf{A}t+\frac{\mathbf{A}^2t^2}{2!}+\frac{\mathbf{A}^3t^3}{3!}+\cdots\right)\mathbf{P}\\ &= \mathbf{P}^{-1}\exp(\mathbf{A}t)\mathbf{P}\end{aligned}$$

6. Write $\exp(\mathbf{A}t) = [\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_n]$, which means $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are the column vectors of $\exp(\mathbf{A}t)$. By Theorem 6.4.7, $\mathbf{x}_k$ is solution to $\mathbf{x}' = \mathbf{Ax}$ for $k = 1, 2, \cdots, n$. Now the Wronskian of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ is

$$W(t) = \det[\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_n] = \det(\exp(\mathbf{A}t))$$

Observe that $W(0) = \det(\exp(\mathbf{0})) = \det(\mathbf{I}) = 1$. Moreover by Abel's theorem for system of differential equations (Theorem 6.1.4), $W(t)$ satisfies the first order linear equation $W'(t) = \mathrm{tr}(\mathbf{A})W(t)$. By solving the initial value problem

$$\begin{cases} W' = \mathrm{tr}(\mathbf{A})W \\ W(0) = 1 \end{cases}$$

we conclude that $\det(\exp(\mathbf{A}t))$ is equal to $W(t) = e^{\mathrm{tr}(\mathbf{A})t}$.

□

The assumption $\mathbf{AB} = \mathbf{BA}$ is necessary in (4) of the above theorem as explained in the following example. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Then $\mathbf{AB} \neq \mathbf{BA}$. Now

$$\exp((\mathbf{A}+\mathbf{B})t) = \exp\left(\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} t\right) = \begin{pmatrix} e^t & e^t-1 \\ 0 & 1 \end{pmatrix}$$

One the other hand

$$\begin{aligned}\exp(\mathbf{A}t) &= \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix}\\ \exp(\mathbf{B}t) &= \mathbf{I}+\mathbf{B}t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\end{aligned}$$

and

$$\exp(\mathbf{A}t)\exp(\mathbf{B}t) = \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^t & te^t \\ 0 & 1 \end{pmatrix} \neq \exp((\mathbf{A}+\mathbf{B})t)$$

Matrix exponential can be used to find the solution of an initial value problem.

Theorem 6.4.9. *The unique solution to the initial value problem*

$$\begin{cases} \mathbf{x}' = \mathbf{Ax} \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases}$$

is

$$\mathbf{x}(t) = \exp(\mathbf{A}t)\mathbf{x}_0$$

Proof. For $\mathbf{x}(t) = \exp(\mathbf{A}t)\mathbf{x}_0$, we have

$$\mathbf{x}'(t) = \frac{d}{dt}\exp(\mathbf{A}t)\mathbf{x}_0 = \mathbf{A}\exp(\mathbf{A}t)\mathbf{x}_0 = \mathbf{A}\mathbf{x}(t)$$

and

$$\mathbf{x}'(0) = \exp(\mathbf{0})\mathbf{x}_0 = \mathbf{I}\mathbf{x}_0 = \mathbf{x}_0$$

Therefore $\mathbf{x}(t) = \exp(\mathbf{A}t)\mathbf{x}_0$ is the solution to the initial value problem. □

Example 6.4.10. *Solve the initial value problem*

$$\begin{cases} \mathbf{x}' = \mathbf{A}\mathbf{x} \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases}$$

where

$$\mathbf{A} = \begin{pmatrix} 5 & 4 \\ -8 & -7 \end{pmatrix} \text{ and } \mathbf{x}_0 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} \lambda - 5 & -4 \\ 8 & \lambda + 7 \end{vmatrix} &= 0 \\ \lambda^2 + 2\lambda - 3 &= 0 \\ \lambda &= 1, -3 \end{aligned}$$

For $\lambda_1 = 1$, an associated eigenvector is

$$\xi_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

For $\lambda_2 = -3$, an associated eigenvector is

$$\xi_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Thus the matrix

$$\mathbf{P} = [\xi_1 \ \xi_2] = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and we have

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix}$$

Hence

$$\begin{aligned} \exp(\mathbf{A}t) &= \mathbf{P}\exp(\mathbf{D}t)\mathbf{P}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} e^t & 0 \\ 0 & e^{-3t} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} e^t & e^{-3t} \\ -e^t & -2e^{-3t} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2e^t - e^{-3t} & e^t - e^{-3t} \\ -2e^t + 2e^{-3t} & -e^t + 2e^{-3t} \end{pmatrix} \end{aligned}$$

Therefore the solution to the initial problem is

$$\begin{aligned}\mathbf{x} &= \exp(\mathbf{A}t)\mathbf{x}_0\\ &= \begin{pmatrix} 2e^t - e^{-3t} & e^t - e^{-3t} \\ -2e^t + 2e^{-3t} & -e^t + 2e^{-3t} \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix}\\ &= \begin{pmatrix} 3e^t - e^{-3t} \\ -3e^t + 2e^{-3t} \end{pmatrix}\end{aligned}$$

□

Exercise 6.4

1. Find $\exp(\mathbf{A}t)$ where $\mathbf{A}$ is the following matrix.
   (a) $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$
   (b) $\begin{pmatrix} 5 & -4 \\ 2 & -1 \end{pmatrix}$
   (c) $\begin{pmatrix} 6 & -6 \\ 4 & -4 \end{pmatrix}$
   (d) $\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$
   (e) $\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$
   (f) $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -8 & -5 & -3 \end{pmatrix}$
   (g) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$
   (h) $\begin{pmatrix} 0 & -4 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$

2. Solve the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ with initial condition $\mathbf{x}(0) = \mathbf{x}_0$ for given $\mathbf{A}$ and $\mathbf{x}_0$.
   (a) $\mathbf{A} = \begin{pmatrix} 2 & 5 \\ -1 & -4 \end{pmatrix}$; $\mathbf{x}_0 = \begin{pmatrix} 1 \\ -5 \end{pmatrix}$
   (b) $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$; $\mathbf{x}_0 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$
   (c) $\mathbf{A} = \begin{pmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{pmatrix}$; $\mathbf{x}_0 = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$
   (d) $\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$; $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$

3. Let $\mathbf{A}$ be a 2×2 matrix. Suppose the eigenvalues of $\mathbf{A}$ are $r = \lambda \pm \mu i$, with $\lambda \in \mathbb{R}$, $\mu > 0$. Let
$$\mathbf{J} = \frac{\mathbf{A} - \lambda\mathbf{I}}{\mu}, \quad \text{where } \mathbf{I} \text{ is the identity matrix.}$$
   (a) Show that $\mathbf{J}^2 = -\mathbf{I}$.
   (b) Show that $\exp(\mathbf{A}t) = e^{\lambda t}(\mathbf{I}\cos\mu t + \mathbf{J}\sin\mu t)$.
   (c) Use the result in (b), or otherwise, to find $\exp(\mathbf{A}t)$ where $\mathbf{A}$ the following matrix.
      (i) $\begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$
      (ii) $\begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$

6.5 Jordan normal forms

When a matrix is not diagonalizable, its matrix exponential can be calculated using Jordan normal form.

Definition 6.5.1. *An $n \times n$ matrix $\mathbf{J}$ is called a* **Jordan matrix** *if it is of the form*

$$\mathbf{J} = \begin{pmatrix} \mathbf{B}_1 & & & \mathit{0} \\ & \mathbf{B}_2 & & \\ & & \ddots & \\ \mathit{0} & & & \mathbf{B}_m \end{pmatrix},$$

where each $\mathbf{B}_i$ is of the form either

$$\lambda_i \mathbf{I} = \begin{pmatrix} \lambda_i & & & \mathit{0} \\ & \lambda_i & & \\ & & \ddots & \\ \mathit{0} & & & \lambda_i \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \lambda_i & 1 & & \mathit{0} \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ \mathit{0} & & & \lambda_i \end{pmatrix}$$

Each $\mathbf{B}_i$ is called a **Jordan block** *of* $\mathbf{J}$.

Note that a diagonal matrix is a Jordan matrix. So Jordan matrix is a generalization of diagonal matrix. The matrix exponential of a diagonal matrix is easy to calculate (Theorem 6.4.2). The matrix exponential of a Jordan matrix can be calculated with a slightly harder effort.

Theorem 6.5.2. *Let*

$$\mathbf{J} = \begin{pmatrix} \mathbf{B}_1 & & & \mathit{0} \\ & \mathbf{B}_2 & & \\ & & \ddots & \\ \mathit{0} & & & \mathbf{B}_m \end{pmatrix}$$

be a Jordan matrix. Then

$$\exp(\mathbf{J}t) = \begin{pmatrix} \exp(\mathbf{B}_1 t) & & & \mathit{0} \\ & \exp(\mathbf{B}_2 t) & & \\ & & \ddots & \\ \mathit{0} & & & \exp(\mathbf{B}_m t) \end{pmatrix}$$

where

$$\exp(\mathbf{B}_i t) = \begin{cases} e^{\lambda_i t} \begin{pmatrix} 1 & & & \mathit{0} \\ & 1 & & \\ & & \ddots & \\ \mathit{0} & & & 1 \end{pmatrix} & \text{if } \mathbf{B}_i = \begin{pmatrix} \lambda_i & & & \mathit{0} \\ & \lambda_i & & \\ & & \ddots & \\ \mathit{0} & & & \lambda_i \end{pmatrix} \\ e^{\lambda_i t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^k}{k!} \\ & 1 & t & \cdots & \frac{t^{k-1}}{(k-1)!} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & t \\ \mathit{0} & & & & 1 \end{pmatrix} & \text{if } \mathbf{B}_i = \begin{pmatrix} \lambda_i & 1 & & \mathit{0} \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ \mathit{0} & & & \lambda_i \end{pmatrix} \end{cases}$$

Proof. Using the property $\exp(\mathbf{A}+\mathbf{B}) = \exp(\mathbf{A})\exp(\mathbf{B})$ if $\mathbf{AB} = \mathbf{BA}$, it suffices to prove the formula for $\exp(\mathbf{B}_i t)$. When $\mathbf{B}_i = \lambda_i \mathbf{I}$, it is obvious that $\exp(\mathbf{I}t) = e^{\lambda_i t}\mathbf{I}$. Finally we have

$$\begin{aligned}
\exp\left[\begin{pmatrix} \lambda_i & 1 & & \text{\Large 0} \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ \text{\Large 0} & & & \lambda_i \end{pmatrix} t\right]
&= \exp\left(\begin{pmatrix} \lambda_i t & 0 & & \text{\Large 0} \\ & \lambda_i t & \ddots & \\ & & \ddots & 0 \\ \text{\Large 0} & & & \lambda_i t \end{pmatrix} + \begin{pmatrix} 0 & t & & \text{\Large 0} \\ & 0 & \ddots & \\ & & \ddots & t \\ \text{\Large 0} & & & 0 \end{pmatrix}\right) \\
&= \exp\begin{pmatrix} \lambda_i t & 0 & & \text{\Large 0} \\ & \lambda_i t & \ddots & \\ & & \ddots & 0 \\ \text{\Large 0} & & & \lambda_i t \end{pmatrix} \exp\begin{pmatrix} 0 & t & & \text{\Large 0} \\ & 0 & \ddots & \\ & & \ddots & t \\ \text{\Large 0} & & & 0 \end{pmatrix} \\
&= e^{\lambda_i t}\begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^n}{n!} \\ & 1 & t & \cdots & \frac{t^{n-1}}{(n-1)!} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & t \\ \text{\Large 0} & & & & 1 \end{pmatrix}.
\end{aligned}$$

The second equality used again the property $\exp(\mathbf{A}+\mathbf{B}) = \exp(\mathbf{A})\exp(\mathbf{B})$ if $\mathbf{AB} = \mathbf{BA}$ and the third equality used the fact that for $k \times k$ matrix

$$\mathbf{N} = \begin{pmatrix} 0 & 1 & & \text{\Large 0} \\ & 0 & \ddots & \\ & & \ddots & 1 \\ \text{\Large 0} & & & 0 \end{pmatrix}$$

we have

$$\mathbf{N}^{k+1} = \mathbf{0}$$

and hence

$$\exp(\mathbf{N}t) = \mathbf{I} + \mathbf{N}t + \frac{1}{2!}\mathbf{N}^2 t^2 + \cdots + \frac{1}{k!}\mathbf{N}^k t^k$$

□

Example 6.5.3. *Find* $\exp(\mathbf{J}t)$ *where* $\mathbf{J}$ *is the following Jordan matrix.*

1. $\begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$
2. $\begin{pmatrix} -3 & 1 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{pmatrix}$
3. $\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$

Solution:

1. $\exp(\mathbf{J}t) = e^{4t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$

2. $\exp(\mathbf{J}t) = e^{-3t}\begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$

3. $\exp(\mathbf{J}t) = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{-t} & t \\ 0 & 0 & e^{-t} \end{pmatrix}$

□

Diagonalizable matrix is a matrix similar to a diagonal matrix. It can be proved that any matrix is similar to a Jordan normal matrix.

Definition 6.5.4 (Jordan normal form)**.** *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. If* $\mathbf{J}$ *is a Jordan matrix similar to* $\mathbf{A}$*, then* $\mathbf{J}$ *is called*[3] *the* **Jordan normal form** *of* $\mathbf{A}$*.*

To diagonalize an $n \times n$ matrix, we need to find n linearly independent eigenvectors. To find the Jordan normal form, we need to find generalized eigenvectors. Recall that a generalized eigenvector of rank k associated with eigenvalue λ is a vector such that

$$\begin{cases} (\mathbf{A} - \lambda\mathbf{I})^{k-1}\eta \neq \mathbf{0}, \\ (\mathbf{A} - \lambda\mathbf{I})^{k}\eta = \mathbf{0}. \end{cases}$$

A vector is a generalized eigenvector of rank 1 if and only if it is an ordinary eigenvector. If η is a generalized eigenvector of rank $k > 1$, then $\eta_i = (\mathbf{A} - \lambda\mathbf{I})^i\eta$ is a generalized eigenvector of rank $k - i$ for $i = 1, 2, \cdots, k - 1$. In particular, $\eta_{k-1} = (\mathbf{A} - \lambda\mathbf{I})^{k-1}\eta$ is an ordinary eigenvector.

Theorem 6.5.5. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. Then there exists non-singular matrix*

$$\mathbf{Q} = [\ \eta_n \quad \eta_{n-1} \quad \cdots \quad \eta_2 \quad \eta_1\ ],$$

where η_i*,* $i = 1, 2, \cdots, n$*, are column vectors of* $\mathbf{Q}$*, such that the following statements hold.*

1. *For any* $i = 1, 2, \cdots, n$*, the vector* η_i *is a generalized eigenvector of* $\mathbf{A}$*.*
2. *If* η_i *is a generalized eigenvector of rank* $k > 1$ *associated with eigenvalue* λ_i*, then*

   $$\eta_{i+1} = (\mathbf{A} - \lambda_i\mathbf{I})\eta_i$$

   In this case η_{i+1} *is a generalized eigenvector of rank* $k - 1$ *associated with the same eigenvalue* λ_i*.*

Furthermore, if $\mathbf{Q}$ *is a non-singular matrix which satisfies the above conditions, then*

$$\mathbf{J} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$$

is the Jordan normal form of $\mathbf{A}$*.*

Note that the Jordan normal form of a matrix is unique up to a permutation of Jordan blocks. We can calculate the matrix exponential of a matrix using its Jordan normal form.

Theorem 6.5.6. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix and* $\mathbf{Q}$ *be a non-singular matrix such that* $\mathbf{J} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ *is the Jordan normal form of* $\mathbf{A}$*. Then*

$$\exp(\mathbf{A}t) = \mathbf{Q}\exp(\mathbf{J}t)\mathbf{Q}^{-1}$$

---

[3] For any square matrix, the Jordan normal form exists and is unique up to permutation of Jordan blocks.

Let's discuss the case for non-diagonalizable 2×2 and 3×3 matrices.

Example 6.5.7. *Let* $\mathbf{A}$ *be a non-diagonalizable* 2×2 *matrix. Then* $\mathbf{A}$ *has only one eigenvalue* λ_1 *and the associated eigenspace is of dimension 1. There exists a generalized eigenvector* η *of rank 2. Let* $\eta_1 = (\mathbf{A} - \lambda_1\mathbf{I})\eta$ *and* $\mathbf{Q} = [\ \eta_1 \quad \eta\ ]$*. The Jordan normal form of* $\mathbf{A}$ *is*

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{J} = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

The minimal polynomial of $\mathbf{A}$ *is* $(x - \lambda_1)^2$*. The matrix exponential is*

$$\exp(\mathbf{A}t) = e^{\lambda_1 t}\mathbf{Q}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\mathbf{Q}^{-1}.$$

Example 6.5.8. *Let* $\mathbf{A}$ *be a non-diagonalizable* 3×3 *matrix. There are 3 possible cases.*

1. *There is one triple eigenvalue* λ_1 *and the associated eigenspace is of dimension 1. Then there exists a generalized eigenvector* η *of rank 3. Let* $\eta_1 = (\mathbf{A} - \lambda_1\mathbf{I})\eta$, $\eta_2 = (\mathbf{A} - \lambda_1\mathbf{I})^2\eta$ *and* $\mathbf{Q} = [\ \eta_2 \quad \eta_1 \quad \eta\ ]$*, we have*

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{J} = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

The minimal polynomial of $\mathbf{A}$ *is* $(x - \lambda_1)^3$*. The matrix exponential is*

$$\exp(\mathbf{A}t) = e^{\lambda_1 t}\mathbf{Q}\begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}\mathbf{Q}^{-1}.$$

2. *There is one triple eigenvalue* λ_1 *and the associated eigenspace is of dimension 2. Then there exists a generalized eigenvector* η *of rank 2 and an eigenvector* ξ *such that* ξ, η *and* $\eta_1 = (\mathbf{A} - \lambda_1\mathbf{I})\eta$ *are linearly independent. Let* $\mathbf{Q} = [\ \xi \quad \eta_1 \quad \eta\ ]$*, we have*

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{J} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

The minimal polynomial of $\mathbf{A}$ *is* $(x - \lambda_1)^2$*. The matrix exponential is*

$$\exp(\mathbf{A}t) = e^{\lambda_1 t}\mathbf{Q}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}\mathbf{Q}^{-1}.$$

3. *There is one simple eigenvalue* λ_1 *and one double eigenvalue* λ_2 *and both of the associated eigenspaces are of dimension 2. Then there exists an eigenvector* ξ *associated with* λ_1 *and a generalized eigenvector* η *of rank 2 associated with* λ_2*. Let* $\eta_1 = (\mathbf{A} - \lambda_2\mathbf{I})\eta$ *(note that* ξ, η, η_1 *must be linearly independent) and* $\mathbf{Q} = [\ \xi \quad \eta_1 \quad \eta\ ]$*, we have*

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{J} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

The minimal polynomial of $\mathbf{A}$ *is* $(x - \lambda_1)(x - \lambda_2)^2$*. The matrix exponential is*

$$\exp(\mathbf{A}t) = \mathbf{Q}\begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & te^{\lambda_2 t} \\ 0 & 0 & e^{\lambda_2 t} \end{pmatrix}\mathbf{Q}^{-1}.$$

Example 6.5.9. *Find* $\exp(\mathbf{A}t)$ *where*

$$\mathbf{A} = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}.$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} \lambda - 5 & 1 \\ -1 & \lambda - 3 \end{vmatrix} &= 0 \\ (\lambda-5)(\lambda-3)+1 &= 0 \\ \lambda^2 - 8\lambda + 16 &= 0 \\ (\lambda-4)^2 &= 0 \\ \lambda &= 4, 4 \end{aligned}$$

we see that $\mathbf{A}$ has only one eigenvalue $\lambda = 4$. Consider

$$\begin{aligned} \det(\mathbf{A} - 4\mathbf{I})\xi &= \mathbf{0} \\ \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\xi &= \mathbf{0} \end{aligned}$$

we can find only one linearly independent eigenvector $\xi = (1, 1)^T$. Thus $\mathbf{A}$ is not diagonalizable. To find $\exp(\mathbf{A}t)$, we need to find a generalized eigenvector of rank 2. Now we take

$$\eta = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and let

$$\begin{cases} \eta_1 = (\mathbf{A} - 4\mathbf{I})\eta = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \eta_2 = (\mathbf{A} - 4\mathbf{I})\eta_1 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{0} \end{cases}$$

We see that η is a generalized eigenvector of rank 2 associated with eigenvalue $\lambda = 4$. We may let

$$\mathbf{Q} = [\eta_1 \ \eta] = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Then

$$\begin{aligned} \mathbf{J} &= \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} \\ &= \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}^{-1}\begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} \end{aligned}$$

is the Jordan normal form of $\mathbf{A}$. Therefore

$$\begin{aligned} \exp(\mathbf{A}t) &= \mathbf{Q}\exp(\mathbf{J}t)\mathbf{Q}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\left(e^{4t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \\ &= e^{4t}\begin{pmatrix} 1+t & -t \\ t & 1-t \end{pmatrix} \end{aligned}$$

□

Example 6.5.10. *Find* $\exp(\mathbf{A}t)$ *where*

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix}.$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} \lambda-3 & -1 & 0 \\ 1 & \lambda-1 & 0 \\ -3 & -2 & \lambda-2 \end{vmatrix} &= 0 \\ (\lambda-2)((\lambda-3)(\lambda-1)+1) &= 0 \\ (\lambda-2)(\lambda^2-4\lambda+4) &= 0 \\ (\lambda-2)^3 &= 0 \\ \lambda &= 2,2,2 \end{aligned}$$

we see that $\mathbf{A}$ has only one eigenvalue $\lambda = 2$. Consider

$$\begin{aligned} \det(\mathbf{A}-2\mathbf{I})\xi &= \mathbf{0} \\ \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 3 & 2 & 0 \end{pmatrix}\xi &= \mathbf{0} \end{aligned}$$

we can find only one linearly independent eigenvector $\xi = (0,0,1)^T$. Thus $\mathbf{A}$ is not diagonalizable. To find $\exp(\mathbf{A}t)$, we need to find a generalized eigenvector of rank 3. Now we take

$$\eta = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and let

$$\begin{cases} \eta_1 = (\mathbf{A}-2\mathbf{I})\eta = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 3 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} \\ \eta_2 = (\mathbf{A}-2\mathbf{I})\eta_1 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 3 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ \eta_3 = (\mathbf{A}-2\mathbf{I})\eta_2 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 3 & 2 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{0} \end{cases}$$

We see that η is a generalized eigenvector of rank 3 associated with eigenvalue $\lambda = 2$. We may let

$$\mathbf{Q} = [\eta_2\ \eta_1\ \eta] = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 3 & 0 \end{pmatrix}$$

Then

$$\begin{aligned}
\mathbf{J} &= \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 3 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 3 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 3 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 3 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}
\end{aligned}$$

is the Jordan normal form of $\mathbf{A}$. Therefore

$$\begin{aligned}
\exp(\mathbf{A}t) &= \mathbf{Q}\exp(\mathbf{J}t)\mathbf{Q}^{-1} \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 3 & 0 \end{pmatrix} \left(e^{2t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} 0 & 3 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \\
&= e^{2t} \begin{pmatrix} 1+t & t & 0 \\ -t & 1-t & 0 \\ 3t+\frac{t^2}{2} & 2t+\frac{t^2}{2} & 1 \end{pmatrix}
\end{aligned}$$

□

Exercise 6.5

1. For the given matrix $\mathbf{A}$, find the Jordan normal form of $\mathbf{A}$ and the matrix exponential $\exp(\mathbf{A}t)$.

   (a) $\begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$

   (b) $\begin{pmatrix} 1 & -4 \\ 4 & -7 \end{pmatrix}$

   (c) $\begin{pmatrix} 5 & -1 & 1 \\ 1 & 3 & 0 \\ -3 & 2 & 1 \end{pmatrix}$

   (d) $\begin{pmatrix} -2 & -9 & 0 \\ 1 & 4 & 0 \\ 1 & 3 & 1 \end{pmatrix}$

   (e) $\begin{pmatrix} -1 & 1 & 1 \\ 1 & 2 & 7 \\ -1 & -3 & -7 \end{pmatrix}$

   (f) $\begin{pmatrix} 2 & 0 & -1 \\ 0 & 4 & 4 \\ 0 & -1 & 0 \end{pmatrix}$

   (g) $\begin{pmatrix} -1 & 3 & -9 \\ 0 & 2 & 0 \\ 1 & -1 & 5 \end{pmatrix}$

   (h) $\begin{pmatrix} 10 & 15 & 22 \\ -4 & -4 & -8 \\ -1 & -3 & -3 \end{pmatrix}$

6.6 Fundamental matrices

Let $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}$ be n linearly independent solutions to the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$. We can put the solutions together and form a matrix $\mathbf{\Psi}(t) = [\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(n)}]$. This matrix is called a fundamental matrix for the system.

Definition 6.6.1 (Fundamental matrix)**.** *A matrix function* $\mathbf{\Psi}(t)$ *is called a* **fundamental matrix** *for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *if the column vectors of* $\mathbf{\Psi}(t)$ *form a fundamental set of solutions for the system.*

We may consider fundamental matrices as solutions to the matrix differential equation $\mathbf{X}'(t) = \mathbf{A}\mathbf{X}(t)$ where $\mathbf{X}(t)$ is an $n \times n$ matrix function of t.

Theorem 6.6.2. *A matrix function* $\mathbf{\Psi}(t)$ *is a fundamental matrix for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *if and only if* $\mathbf{\Psi}(t)$ *satisfies the matrix differential equation* $\frac{d\mathbf{\Psi}}{dt} = \mathbf{A}\mathbf{\Psi}$ *and* $\mathbf{\Psi}(t_0)$ *is non-singular for some* t_0*.*

Proof. For any vector valued functions $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \cdots, \mathbf{x}^{(n)}(t)$, consider the matrix

$$\mathbf{\Psi}(t) = \begin{bmatrix} \mathbf{x}^{(1)} & \mathbf{x}^{(2)} & \cdots & \mathbf{x}^{(n)} \end{bmatrix}.$$

We have

$$\frac{d\mathbf{\Psi}}{dt} = \begin{bmatrix} \frac{d\mathbf{x}^{(1)}}{dt} & \frac{d\mathbf{x}^{(2)}}{dt} & \cdots & \frac{d\mathbf{x}^{(n)}}{dt} \end{bmatrix}$$

and

$$\mathbf{A}\mathbf{\Psi} = \begin{bmatrix} \mathbf{A}\mathbf{x}^{(1)} & \mathbf{A}\mathbf{x}^{(2)} & \cdots & \mathbf{A}\mathbf{x}^{(n)} \end{bmatrix}.$$

Thus $\mathbf{\Psi}$ satisfies the equation

$$\frac{d\mathbf{\Psi}}{dt} = \mathbf{A}\mathbf{\Psi}$$

if and only if $\mathbf{x}^{(i)}$ is a solution to the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for any $i = 1, 2, \cdots, n$. Now the solutions $\mathbf{x}^{(1)}(t), \mathbf{x}^{(2)}(t), \cdots, \mathbf{x}^{(n)}(t)$ form a fundamental set of solutions if and only if they are linearly independent at some t_0 if and only if $\mathbf{\Psi}(t_0)$ is non-singular for some t_0. □

Theorem 6.6.3. *The matrix exponential* $\exp(\mathbf{A}t)$ *is a fundamental matrix for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$.

Proof. The matrix exponential $\exp(\mathbf{A}t)$ satisfies $\frac{d}{dt}\exp(\mathbf{A}t) = \mathbf{A}\exp(\mathbf{A}t)$ and when $t = 0$, the value of $\exp(\mathbf{A}t)$ is $\mathbf{I}$ which is non-singular. Thus $\exp(\mathbf{A}t)$ is a fundamental matrix for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ by Theorem 6.6.2. □

If $\mathbf{\Psi}$ is a fundamental matrix, we can multiply any non-singular matrix from the right to obtain another fundamental matrix.

Theorem 6.6.4. *Let* $\mathbf{\Psi}$ *be a fundamental matrix for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *and* $\mathbf{P}$ *be a non-singular constant matrix. Then* $\mathbf{\Psi}(t)\mathbf{P}$ *is also a fundamental matrix for the system.*

Proof. Observe that

$$\frac{d}{dt}(\mathbf{\Psi}(t)\mathbf{P}) = \mathbf{\Psi}'(t)\mathbf{P} = \mathbf{A}\mathbf{\Psi}(t)\mathbf{P}$$

and $\mathbf{\Psi}(t)\mathbf{P}$ is non-singular for any t. Thus $\mathbf{\Psi}(t)\mathbf{P}$ is a fundamental matrix for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$. □

Caution: In general $\mathbf{P}\mathbf{\Psi}(t)$ is not a fundamental matrix.

The theorem below is useful in finding a fundamental matrix for a system.

Theorem 6.6.5. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix.*

1. *Suppose* $\mathbf{P}$ *is a non-singular matrix such that* $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ *is a diagonal matrix. Then*

$$\mathbf{\Psi}(t) = \mathbf{P}\exp(\mathbf{D}t)$$

is a fundamental matrix for the system $\mathbf{x}' = \mathbf{Ax}$*.*

2. *Suppose* $\mathbf{Q}$ *is a non-singular matrix such that* $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$ *is a Jordan matrix. Then*

$$\mathbf{\Psi}(t) = \mathbf{Q}\exp(\mathbf{J}t)$$

is a fundamental matrix for the system $\mathbf{x}' = \mathbf{Ax}$*.*

Proof. For the first statement, since $\exp(\mathbf{A}t) = \mathbf{P}\exp(\mathbf{D}t)\mathbf{P}^{-1}$ is a fundamental matrix (Theorem 6.6.3) for the system $\mathbf{x}' = \mathbf{Ax}$, we have $\mathbf{\Psi}(t) = \mathbf{P}\exp(\mathbf{D}t) = \exp(\mathbf{A}t)\mathbf{P}$ is also a fundamental matrix by Theorem 6.6.4. The proof of the second statement is similar. □

Example 6.6.6. *Find a fundamental matrix for the system*

$$\mathbf{x}' = \begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix}\mathbf{x}.$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} \lambda - 2 & -2 \\ -3 & \lambda - 1 \end{vmatrix} &= 0 \\ \lambda^2 - 3\lambda - 4 &= 0 \\ \lambda &= 4, -1. \end{aligned}$$

For $\lambda_1 = 4$, an associated eigenvector is

$$\xi_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For $\lambda_2 = -1$, an associated eigenvector is

$$\xi_2 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

Hence the matrix

$$\mathbf{P} = [\xi_1\ \xi_2] = \begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and we have

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}.$$

Therefore a fundamental matrix for the system is

$$\begin{aligned} \mathbf{\Psi}(t) &= \mathbf{P}\exp(\mathbf{D}t) \\ &= \begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} e^{4t} & 0 \\ 0 & e^{-t} \end{pmatrix} \\ &= \begin{pmatrix} e^{4t} & 2e^{-t} \\ e^{4t} & -3e^{-t} \end{pmatrix} \end{aligned}$$

□

Example 6.6.7. *Find a fundamental matrix for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *where*

$$\mathbf{A} = \begin{pmatrix} 1 & -3 \\ 3 & 7 \end{pmatrix}.$$

Solution: By solving the characteristic equation, we see that $\mathbf{A}$ has only one eigenvalue $\lambda = 4$. Taking $\eta = (1, 0)^T$, we have

$$\begin{cases} \eta_1 = (\mathbf{A} - 4\mathbf{I})\eta = \begin{pmatrix} -3 & -3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix} \\ \eta_2 = (\mathbf{A} - 4\mathbf{I})\eta_1 = \begin{pmatrix} -3 & -3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} -3 \\ 3 \end{pmatrix} = \mathbf{0} \end{cases}$$

(In fact $\eta_2 = \mathbf{0}$ is automatic by Cayley-Hamilton theorem since $\mathbf{A}$ has only one eigenvalue $\lambda = 4$.) Thus η is a generalized eigenvector of rank 2 associated with $\lambda = 4$. Now we take

$$\mathbf{Q} = [\eta_1 \ \eta] = \begin{pmatrix} -3 & 1 \\ 3 & 0 \end{pmatrix}$$

Then

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} = \mathbf{J}$$

is the Jordan normal form of $\mathbf{A}$. Therefore a fundamental matrix for the system is

$$\begin{aligned} \mathbf{\Psi}(t) &= \mathbf{Q}\exp(\mathbf{J}t) \\ &= \begin{pmatrix} -3 & 1 \\ 3 & 0 \end{pmatrix} \left(e^{4t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \right) \\ &= e^{4t} \begin{pmatrix} -3 & 1 - 3t \\ 3 & 3t \end{pmatrix} \end{aligned}$$

□

Example 6.6.8. *Find a fundamental matrix for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *where*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 \\ -5 & -3 & -7 \\ 1 & 0 & 0 \end{pmatrix}.$$

Solution: By solving the characteristic equation, we see that $\mathbf{A}$ has only one eigenvalue $\lambda = -1$. Taking $\eta = (1, 0, 0)^T$, we have

$$\begin{cases} \eta_1 = (\mathbf{A} + \mathbf{I})\eta = \begin{pmatrix} 1 & 1 & 2 \\ -5 & -2 & -7 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} \\ \eta_2 = (\mathbf{A} + \mathbf{I})\eta_1 = \begin{pmatrix} 1 & 1 & 2 \\ -5 & -2 & -7 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -5 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix} \\ \eta_3 = (\mathbf{A} + \mathbf{I})\eta_2 = \begin{pmatrix} 1 & 1 & 2 \\ -5 & -2 & -7 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ -2 \\ 2 \end{pmatrix} = \mathbf{0} \end{cases}$$

(In fact $\eta_3 = \mathbf{0}$ is automatic by Cayley-Hamilton theorem since $\mathbf{A}$ has only one eigenvalue $\lambda = -1$.) Thus η is a generalized eigenvector of rank 3 associated with $\lambda = -1$. Now we take

$$\mathbf{Q} = [\ \eta_2 \quad \eta_1 \quad \eta\ ] = \begin{pmatrix} -2 & 1 & 1 \\ -2 & -5 & 0 \\ 2 & 1 & 0 \end{pmatrix}$$

Then

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} = \mathbf{J}$$

is the Jordan form of $\mathbf{A}$. Therefore a fundamental matrix for the system is

$$\begin{aligned} \mathbf{\Psi} &= \mathbf{Q}\exp(\mathbf{J}t) \\ &= \begin{pmatrix} -2 & 1 & 1 \\ -2 & -5 & 0 \\ 2 & 1 & 0 \end{pmatrix} \left(e^{-t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &= e^{-t} \begin{pmatrix} -2 & 1-2t & 1+t-t^2 \\ -2 & -5-t & -5t-t^2 \\ 2 & 1-2t & t+t^2 \end{pmatrix} \end{aligned}$$

□

We have seen how $\exp(\mathbf{A}t)$ can be used to write down the solution of an initial value problem (Theorem 6.4.9). We can also use it to find fundamental matrix with initial condition.

Theorem 6.6.9. *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix. For any* $n \times n$ *non-singular matrix* $\mathbf{\Psi}_0$*, the unique fundamental matrix* $\mathbf{\Psi}(t)$ *for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *which satisfies the initial condition* $\mathbf{\Psi}(t) = \mathbf{\Psi}_0$ *is* $\mathbf{\Psi}(t) = \exp(\mathbf{A}t)\mathbf{\Psi}_0$.

Proof. For $\mathbf{\Psi}(t) = \exp(\mathbf{A}t)\mathbf{\Psi}_0$, we have

$$\mathbf{\Psi}'(t) = \frac{d}{dt}(\exp(\mathbf{A}t)\mathbf{\Psi}_0) = \mathbf{A}\exp(\mathbf{A}t)\mathbf{\Psi}_0 = \mathbf{A}\mathbf{\Psi}(t)$$

Moreover $\mathbf{\Psi}(0) = \exp(\mathbf{0})\mathbf{\Psi}_0 = \mathbf{I}\mathbf{\Psi}_0 = \mathbf{\Psi}_0$ and is non-singular. Therefore $\mathbf{\Psi}(t)$ is a fundamental matrix with $\mathbf{\Psi}(0) = \mathbf{\Psi}_0$. Such fundamental matrix is unique by the uniqueness of solution to initial value problem (Theorem 6.1.2). □

Example 6.6.10. *Find a fundamental matrix* $\mathbf{\Psi}(t)$ *for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x}$ *with* $\mathbf{\Psi}(0) = \mathbf{\Psi}_0$ *where*

$$\mathbf{A} = \begin{pmatrix} 1 & 9 \\ -1 & -5 \end{pmatrix} \text{ and } \mathbf{\Psi}_0 = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$$

Solution: By solving the characteristic equation, $\mathbf{A}$ has only one eigenvalue $\lambda = -2$. Taking $\eta = (1, 0)^T$, we have

$$\eta_1 = (\mathbf{A} + 2\mathbf{I})\eta = \begin{pmatrix} 3 & 9 \\ -1 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

and $\eta_2 = (\mathbf{A} + 2\mathbf{I})^2\eta = \mathbf{0}$ by Cayley-Hamilton theorem. Thus η is a generalized eigenvector associated with $\lambda = -2$. Now taking

$$\mathbf{Q} = [\eta_1\ \eta] = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix}$$

we have

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix} = \mathbf{J}$$

is the Jordan normal form of $\mathbf{A}$. Thus

$$\begin{aligned} \exp(\mathbf{A}t) &= \mathbf{Q}\exp(\mathbf{J}t)\mathbf{Q}^{-1} \\ &= \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix}\left(e^{-2t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \\ &= e^{-2t}\begin{pmatrix} 1+3t & 9t \\ -t & 1-3t \end{pmatrix} \end{aligned}$$

Therefore the required fundamental matrix with initial condition is

$$\begin{aligned} \mathbf{\Psi}(t) &= \exp(\mathbf{A}t)\mathbf{\Psi}_0 \\ &= e^{-2t}\begin{pmatrix} 1+3t & 9t \\ -t & 1-3t \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \\ &= e^{-2t}\begin{pmatrix} 1-6t & 18t \\ -1+2t & 2-6t \end{pmatrix} \end{aligned}$$

Exercise 6.6

1. Find a fundamental matrix for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is the following matrix.

   (a) $\begin{pmatrix} 3 & -2 \\ 2 & -2 \end{pmatrix}$

   (b) $\begin{pmatrix} 1 & 1 \\ 4 & -2 \end{pmatrix}$

   (c) $\begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$

   (d) $\begin{pmatrix} 3 & 4 \\ -1 & -2 \end{pmatrix}$

   (e) $\begin{pmatrix} -1 & -4 \\ 1 & -1 \end{pmatrix}$

   (f) $\begin{pmatrix} 1 & -3 \\ 3 & -5 \end{pmatrix}$

   (g) $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -8 & -5 & -3 \end{pmatrix}$

   (h) $\begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$

   (i) $\begin{pmatrix} 3 & 0 & 0 \\ -4 & 7 & -4 \\ -2 & 2 & 1 \end{pmatrix}$

   (j) $\begin{pmatrix} 3 & 1 & 3 \\ 2 & 2 & 2 \\ -1 & 0 & 1 \end{pmatrix}$

   (k) $\begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$

   (l) $\begin{pmatrix} -2 & -2 & 3 \\ 0 & -1 & -1 \\ 0 & 1 & -3 \end{pmatrix}$

2. Find the fundamental matrix $\mathbf{\Phi}$ which satisfies $\mathbf{\Phi}(0) = \mathbf{\Phi}_0$ for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for the given matrices $\mathbf{A}$ and $\mathbf{\Phi}_0$.

   (a) $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ -1 & -2 \end{pmatrix}$; $\mathbf{\Phi}_0 = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}$

   (b) $\mathbf{A} = \begin{pmatrix} 7 & 1 \\ -4 & 3 \end{pmatrix}$; $\mathbf{\Phi}_0 = \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}$

   (c) $\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 \\ -4 & 7 & -4 \\ -2 & 2 & 1 \end{pmatrix}$; $\mathbf{\Phi}_0 = \begin{pmatrix} 2 & 0 & -1 \\ 0 & -3 & 1 \\ -1 & 1 & 0 \end{pmatrix}$

(d) $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -3 & 2 & 4 \end{pmatrix}$; $\mathbf{\Phi}_0 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$

3. Let $\mathbf{A}$ be a square matrix and $\mathbf{\Psi}$ be a fundamental matrix for the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$.
   (a) Prove that for any non-singular constant matrix $\mathbf{Q}$, we have $\mathbf{Q}\mathbf{\Psi}$ is a fundamental matrix for the system if and only if $\mathbf{QA} = \mathbf{AQ}$.
   (b) Prove that $(\mathbf{\Psi}^T)^{-1}$ is a fundamental matrix for the system $\mathbf{x}' = -\mathbf{A}^T\mathbf{x}$.
4. Prove that if $\mathbf{\Psi}_1(t)$ and $\mathbf{\Psi}_2(t)$ are two fundamental matrices for a system, then $\mathbf{\Psi}_2 = \mathbf{\Psi}_1\mathbf{P}$ for some non-singular matrix $\mathbf{P}$.

6.7 Nonhomogeneous linear systems

We now turn to nonhomogeneous system

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{g}(t)$$

where $\mathbf{g}(t)$ is a continuous vector valued function. The general solution of the system can be expressed as

$$\mathbf{x} = c_1\mathbf{x}^{(1)} + \cdots + c_n\mathbf{x}^{(n)} + \mathbf{x}_p(t),$$

where $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)}$ is a fundamental set of solutions to the associated homogeneous system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ and $\mathbf{x}_p$ is a particular solution to the nonhomogeneous system. So to solve the nonhomogeneous system, it suffices to find a particular solution. We will briefly describe two methods for finding a particular solution.

Variation of parameters
The first method we introduce is variation of parameters.

Theorem 6.7.1. *Let $\mathbf{\Psi}(t)$ be a fundamental matrix for the system $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ and $\mathbf{g}(t)$ be a continuous vector valued function. Then a particular solution to the nonhomogeneous system*

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{g}(t)$$

is given by

$$\mathbf{x}_p = \mathbf{\Psi}(t)\int \mathbf{\Psi}^{-1}(t)\mathbf{g}(t)dt$$

Moreover the solution to the initial value problem

$$\begin{cases} \mathbf{x}'(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{g}(t) \\ \mathbf{x}(t_0) &= \mathbf{x}_0 \end{cases}$$

is given by

$$\mathbf{x}(t) = \mathbf{\Psi}(t)\left(\mathbf{\Psi}^{-1}(t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{\Psi}^{-1}(s)\mathbf{g}(s)ds\right)$$

In particular the solution to the nonhomogeneous system with initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is

$$\mathbf{x}(t) = \exp(\mathbf{A}t)\left(\mathbf{x}_0 + \int_0^t \exp(-\mathbf{A}s)\mathbf{g}(s)ds\right)$$

Proof. We check that $\mathbf{x}_p = \mathbf{\Psi}(t)\int \mathbf{\Psi}^{-1}(t)\mathbf{g}(t)dt$ satisfies the nonhomogeneous system.

$$\begin{aligned}
\mathbf{x}_p' &= \frac{d}{dx}\left(\mathbf{\Psi}(t)\int \mathbf{\Psi}^{-1}(t)\mathbf{g}(t)dt\right)\\
&= \mathbf{\Psi}'\int \mathbf{\Psi}^{-1}\mathbf{g}dt + \mathbf{\Psi}\frac{d}{dt}\left(\int \mathbf{\Psi}^{-1}\mathbf{g}dt\right)\\
&= \mathbf{A\Psi}\int \mathbf{\Psi}^{-1}\mathbf{g}dt + \mathbf{\Psi\Psi}^{-1}\mathbf{g}\\
&= \mathbf{A}\mathbf{x}_p + \mathbf{g}
\end{aligned}$$

Now

$$\mathbf{x}(t) = \mathbf{\Psi}(t)\left(\mathbf{\Psi}^{-1}(t_0)\mathbf{x}_0 + \int_{t_0}^{t}\mathbf{\Psi}^{-1}(s)\mathbf{g}(s)ds\right)$$

satisfies the nonhomogeneous system. Since $\mathbf{x}$ satisfies the initial condition

$$\begin{aligned}
\mathbf{x}(t_0) &= \mathbf{\Psi}(t_0)\left(\mathbf{\Psi}^{-1}(t_0)\mathbf{x}_0 + \int_{t_0}^{t_0}\mathbf{\Psi}^{-1}(s)\mathbf{g}(s)ds\right)\\
&= \mathbf{\Psi}(t_0)\mathbf{\Psi}^{-1}(t_0)\mathbf{x}_0\\
&= \mathbf{x}_0
\end{aligned}$$

it is the unique solution to the initial value problem. In particular, $\exp(\mathbf{A}t)$ is a fundamental matrix which is equal to the identity matrix $\mathbf{I}$ when $t = 0$. Therefore the solution to the nonhomogeneous system with initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is

$$\begin{aligned}
\mathbf{x} &= \exp(\mathbf{A}t)\left((\exp(\mathbf{0}))^{-1}\mathbf{x}_0 + \int_0^t (\exp(\mathbf{A}s))^{-1}\mathbf{g}(s)ds\right)\\
&= \exp(\mathbf{A}t)\left(\mathbf{x}_0 + \int_0^t \exp(-\mathbf{A}s)\mathbf{g}(s)ds\right)
\end{aligned}$$

Here we used the fact that the inverse of $\exp(\mathbf{A}s)$ is $\exp(-\mathbf{A}s)$. □

The above theorem works even when the coefficient matrix of the system is not constant. Suppose $\mathbf{\Psi}(t)$ is a fundamental matrix for the homogeneous system $\mathbf{x}'(t) = \mathbf{P}(t)\mathbf{x}(t)$ where $\mathbf{P}(t)$ is a continuous matrix valued function. That means $\mathbf{\Psi}'(t) = \mathbf{P}(t)\mathbf{\Psi}(t)$ and $\mathbf{\Psi}(t_0)$ is non-singular for some t_0. Then a particular solution to the nonhomogeneous system $\mathbf{x}'(t) = \mathbf{P}(t)\mathbf{x}(t) + \mathbf{g}(t)$ is

$$\mathbf{x}_p(t) = \mathbf{\Psi}(t)\int \mathbf{\Psi}^{-1}(t)\mathbf{g}(t)dt$$

and the solution to the system with initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ is

$$\mathbf{x}(t) = \mathbf{\Psi}(t)\left(\mathbf{\Psi}^{-1}(t_0)\mathbf{x}_0 + \int_{t_0}^{t}\mathbf{\Psi}^{-1}(s)\mathbf{g}(s)ds\right)$$

Example 6.7.2. *Use method of variation of parameters to find a particular solution for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{g}(t)$ *where*

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \quad \textit{and} \quad \mathbf{g}(t) = \begin{pmatrix} 4e^{-t} \\ 18t \end{pmatrix}.$$

Solution: The eigenvalues of $\mathbf{A}$ are $\lambda_1 = -3$ and $\lambda_2 = -1$ with eigenvectors $(1, -1)^T$ and $(1, 1)^T$ respectively. Thus a fundamental matrix of the system is

$$\mathbf{\Psi} = \mathbf{P}\exp(\mathbf{D}t) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e^{-3t} & 0 \\ 0 & e^{-t} \end{pmatrix} = \begin{pmatrix} e^{-3t} & e^{-t} \\ -e^{-3t} & e^{-t} \end{pmatrix}.$$

Now

$$\begin{aligned} \mathbf{\Psi}^{-1}\mathbf{g} &= \begin{pmatrix} e^{-3t} & e^{-t} \\ -e^{-3t} & e^{-t} \end{pmatrix}^{-1}\begin{pmatrix} 4e^{-t} \\ 18t \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} e^{3t} & -e^{3t} \\ e^{t} & e^{t} \end{pmatrix}\begin{pmatrix} 4e^{-t} \\ 18t \end{pmatrix} \\ &= \begin{pmatrix} 2e^{2t} - 9te^{3t} \\ 2 + 9te^{t} \end{pmatrix}. \end{aligned}$$

Thus

$$\begin{aligned} \int \mathbf{\Psi}^{-1}\mathbf{g}dt &= \int \begin{pmatrix} 2e^{2t} - 9te^{3t} \\ 2 + 9te^{t} \end{pmatrix} dt \\ &= \begin{pmatrix} e^{2t} - 3te^{3t} + e^{3t} + c_1 \\ 2t + 9te^{t} - 9e^{t} + c_2 \end{pmatrix}. \end{aligned}$$

Therefore a particular solution is

$$\begin{aligned} \mathbf{x}_p &= \begin{pmatrix} e^{-3t} & e^{-t} \\ -e^{-3t} & e^{-t} \end{pmatrix}\begin{pmatrix} e^{2t} - 3te^{3t} + e^{3t} \\ 2t + 9te^{t} - 9e^{t} \end{pmatrix} \\ &= \begin{pmatrix} e^{-3t}(e^{2t} - 3te^{3t} + e^{3t}) + e^{-t}(2t + 9te^{t} - 9e^{t}) \\ -e^{-3t}(e^{2t} - 3te^{3t} + e^{3t}) + e^{-t}(2t + 9te^{t} - 9e^{t}) \end{pmatrix} \\ &= \begin{pmatrix} 2te^{-t} + e^{-t} + 6t - 8 \\ 2te^{-t} - e^{-t} + 12t - 10 \end{pmatrix} \end{aligned}$$

□

Example 6.7.3. *Solve the initial value problem* $\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{g}(t)$ *with* $\mathbf{x}(0) = (-1, 1)^T$*, where*

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad \mathbf{g}(t) = \begin{pmatrix} 2e^{-t} \\ 0 \end{pmatrix}.$$

Solution: Observe that $\mathbf{A}$ is a Jordan matrix. We have

$$\exp(\mathbf{A}t) = e^{t}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

Then we calculate

$$\begin{aligned} \exp(-\mathbf{A}s) &= e^{-s}\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} e^{-s} & -se^{-s} \\ 0 & e^{-s} \end{pmatrix} \\ \int_0^t \exp(-\mathbf{A}s)\mathbf{g}(s)ds &= \int_0^t \begin{pmatrix} e^{-s} & -se^{-s} \\ 0 & e^{-s} \end{pmatrix}\begin{pmatrix} 2e^{-s} \\ 0 \end{pmatrix} ds \\ &= \int_0^t \begin{pmatrix} 2e^{-2s} \\ 0 \end{pmatrix} ds \\ &= \begin{pmatrix} 1 - e^{-2t} \\ 0 \end{pmatrix} \end{aligned}$$

Therefore the solution to the initial value problem is

$$\begin{aligned}
\mathbf{x}_p &= \exp(\mathbf{A}t)\left(\mathbf{x}_0 + \int_0^t \exp(-\mathbf{A}s)\mathbf{g}(s)ds\right)\\
&= e^t\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1-e^{-2t} \\ 0 \end{pmatrix}\right)\\
&= e^t\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -e^{-2t} \\ 1 \end{pmatrix}\\
&= \begin{pmatrix} te^t - e^{-t} \\ e^t \end{pmatrix}
\end{aligned}$$

Undetermined coefficients We are not going to discuss the general case of the method of undetermined coefficients. We only study the nonhomogeneous system

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) + t^k e^{\alpha t}\mathbf{g}$$

where $\mathbf{g}$ is a constant vector. Then there is a particular solution of the form

$$\mathbf{x}_p = e^{\alpha t}(t^{m+k}\mathbf{a}_{m+k} + t^{m+k-1}\mathbf{a}_{m+k-1} + \cdots + t\mathbf{a}_1 + \mathbf{a}_0)$$

where m is the smallest non-negative integer such that the general solution of the associated homogeneous system does not contain any term of the form $t^m e^{\alpha t}\mathbf{a}$ and $\mathbf{a}_{m+k}, \mathbf{a}_{m+k-1}, \cdots, \mathbf{a}_1, \mathbf{a}_0$ are constant vectors which can be determined by substituting $\mathbf{x}_p$ to the system. Note that a particular solution may contain a term of the form $t^i e^{\alpha t}\mathbf{a}_i$ even if it appears in the general solution of the associated homogeneous system.

Example 6.7.4. *Use method of undetermined coefficients to find a particular solution for the system* $\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{g}(t)$ *where*

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \quad \textit{and} \quad \mathbf{g}(t) = \begin{pmatrix} 4e^{-t} \\ 18t \end{pmatrix}.$$

Solution: Let

$$\mathbf{x}_p(t) = te^{-t}\mathbf{a} + e^{-t}\mathbf{b} + t\mathbf{c} + \mathbf{d}$$

be a particular solution. (Remark: It is not surprising that the term $te^{-t}\mathbf{a}$ appears since $\lambda = -1$ is an eigenvalue of $\mathbf{A}$. But one should note that we also need the term $e^{-t}\mathbf{b}$. It is because $\mathbf{b}$ may not be an eigenvector. Thus $e^{-t}\mathbf{b}$ may not be a solution to the associated homogeneous system and may appear in the particular solution.) Substituting $\mathbf{x}_p$ to the nonhomogeneous system, we have

$$\begin{aligned}
\mathbf{x}_p' &= \mathbf{A}\mathbf{x}_p + \mathbf{g}(t)\\
-te^{-t}\mathbf{a} + e^{-t}(\mathbf{a}-\mathbf{b}) + \mathbf{c} &= te^{-t}\mathbf{A}\mathbf{a} + e^{-t}\mathbf{A}\mathbf{b} + t\mathbf{A}\mathbf{c} + \mathbf{A}\mathbf{d} + \begin{pmatrix} 4e^{-t} \\ 18t \end{pmatrix}
\end{aligned}$$

Comparing the coefficients of $te^{-t}, e^{-t}, t, 1$, we have

$$\begin{cases}
(\mathbf{A}+\mathbf{I})\mathbf{a} = \mathbf{0}\\
(\mathbf{A}+\mathbf{I})\mathbf{b} = \mathbf{a} - \begin{pmatrix} 4 \\ 0 \end{pmatrix}\\
\mathbf{A}\mathbf{c} = -\begin{pmatrix} 0 \\ 18 \end{pmatrix}\\
\mathbf{A}\mathbf{d} = \mathbf{c}
\end{cases}$$

The solution for $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ of the above equations is not unique and we may take any solution, say

$$\mathbf{a} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 6 \\ 12 \end{pmatrix}, \quad \mathbf{d} = -\begin{pmatrix} 8 \\ 10 \end{pmatrix}.$$

to obtain a particular solution

$$\begin{aligned} \mathbf{x}_p &= te^{-t}\begin{pmatrix} 2 \\ 2 \end{pmatrix} - e^{-t}\begin{pmatrix} 0 \\ 2 \end{pmatrix} + t\begin{pmatrix} 6 \\ 12 \end{pmatrix} - \begin{pmatrix} 8 \\ 10 \end{pmatrix} \\ &= \begin{pmatrix} 2te^{-t} + 6t - 8 \\ 2te^{-t} - 2e^{-t} + 12t - 10 \end{pmatrix} \end{aligned}$$

(Note: This particular solution is not the same as the one given in Example 6.7.2. They are different by $(e^{-t}, e^{-t})^T$ which is a solution to the associated homogeneous system.) ☐

Exercise 6.7

1. Use the method of variation of parameters to find a particular solution for each of the following non-homogeneous equations.

   (a) $\mathbf{x}' = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\mathbf{x} + \begin{pmatrix} -6e^{5t} \\ 6e^{5t} \end{pmatrix}$

   (b) $\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ 4 & -2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} e^{-2t} \\ -2e^{t} \end{pmatrix}$

   (c) $\mathbf{x}' = \begin{pmatrix} 2 & -1 \\ 4 & -3 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 0 \\ 9e^{t} \end{pmatrix}$

   (d) $\mathbf{x}' = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}\mathbf{x} + \begin{pmatrix} e^{t} \\ 0 \end{pmatrix}$

   (e) $\mathbf{x}' = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} e^{t} \\ t \end{pmatrix}$

   (f) $\mathbf{x}' = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 0 \\ \cos t \end{pmatrix}$

2. For each of the nonhomogeneous linear systems in Question 1, write down a suitable form $\mathbf{x}_p(t)$ of a particular solution.

7 Answers to exercises

Exercise 1.1

1. (a) $y = e^{3x} + Ce^{-x}$
   (b) $y = 3x + Cx^{-\frac{1}{3}}$
   (c) $y = \frac{1}{3} + Ce^{-x^3}$
   (d) $y = \frac{C + \ln|x|}{x}$
   (e) $y = \frac{2}{3}\sqrt{x} + \frac{C}{x}$
   (f) $y = Cx - x\cos x$
   (g) $y = \frac{2}{3}(x+1)^{\frac{7}{2}} + C(x+1)^2$
   (h) $y = \sin x + C\cos x$
   (i) $y = (1 + \frac{C}{x})e^{-3x}$

2. (a) $y = e^{2x}$
   (b) $y = 1 + e^{-\sin x}$
   (c) $y = 1 + 16(x^2+4)^{-\frac{3}{2}}$
   (d) $y = \frac{x\ln x - x + 21}{x+1}$
   (e) $y = \frac{\ln x}{x} - \frac{1}{x} + \frac{3}{x^2}$
   (f) $y = \frac{\pi - 1 - \cos x}{x}$

Exercise 1.2

1. (a) $y = \frac{1}{x^2 + C}$
   (b) $y = (x^{\frac{3}{2}} + C)^2$
   (c) $y = 1 + (x^2 + C)^3$
   (d) $y = Ce^{-\cos x}$
   (e) $y^2 + 1 = Ce^{x^2}$
   (f) $\ln|1+y| = x + \frac{1}{2}x^2 + C$

2. (a) $y = xe^{x^2-1}$
   (b) $y = 2e^{e^x}$
   (c) $y^2 = 1 + \sqrt{x^2 - 16}$
   (d) $y = -3e^{x^4 - x}$
   (e) $y = \frac{\pi}{2}\sin x$
   (f) $y = \tan\left(x^3 + \frac{\pi}{4}\right)$

3. $y = \frac{1000}{1 + 9e^{-0.08x}}$

Exercise 1.3

1. (a) $\frac{5}{2}x^2 + 4xy - 2y^4 = C$
   (b) $x^3 + 2y^2x + 2y^3 = C$
   (c) $\frac{3}{2}x^2y^2 - xy^3 = C$
   (d) $x + e^{xy} + y^2 = C$
   (e) $3x^4 + 4y^3 + 12y\ln x = C$
   (f) $\sin x + x\ln y + e^y = C$

2. (a) $k = 2$; $x^2y^2 - 3x + 4y = C$
   (b) $k = -3$; $3x^2y - xy^3 + 2y^2 = C$
   (c) $k = 2$; $x^3 + x^2y^2 + y^4 = C$
   (d) $k = 4$; $5x^3y^3 + 5xy^4 = C$

3. (a) $x^3y + \frac{1}{2}x^2y^2 = C$
   (b) $\frac{x}{y} = C$
   (c) $\ln|xy| + \frac{y^3}{3} = C$
   (d) $\frac{\ln(x^2+y^2)}{2} + \tan^{-1}\frac{y}{x}$

Exercise 1.4

1. (a) $y^2 = x^2(\ln|x| + C)$
   (b) $y = x\sin(C + \ln|x|)$
   (c) $y = x(C + \ln|x|)^2$
   (d) $\ln|xy| = xy^{-1} + C$
   (e) $y = \frac{x}{C - \ln|x|}$
   (f) $y = -x\ln(C - \ln|x|)$

Exercise 1.5

1.

(a) $y = \frac{1}{Cx-x^2}$ (b) $y^3 = \frac{7x}{15+Cx^7}$ (c) $y = \frac{1}{Cx-x^2}$

Exercise 1.6

1. (a) $y = e^{x^2+\frac{C}{x^2}}$
   (b) $x = 2(x+y)^{\frac{1}{2}} - 2\ln(1+(x+y)^{\frac{1}{2}}) + C$
   (c) $y = \tan(x+C) - x - 3$
   (d) $y = -\ln(Ce^x - 1)$
2. (a) $y = x + \frac{2x}{Ce^{\frac{2}{x}}-1}$
   (b) $y = \frac{1}{x} + \frac{3x^2}{C-x^3}$

Exercise 1.7

1. (a) $x = C_1y^2 + C_2$
   (b) $y = C_1\cos 2x + C_2\sin 2x$
   (c) $y = x^2 + C_1\ln x + C_2$
   (d) $y = \ln x + C_1x^{-2} + C_2$
   (e) $y = \pm(C_1 + C_2e^x)^{\frac{1}{2}}$
   (f) $y^3 + C_1y + 3x + C_2 = 0$
2. (a) $y = \frac{1}{\sqrt{C-x^2}}$ (Separable)
   (b) $y = x^2\ln|x| + Cx^2$ (Linear)
   (c) $3x^3 + xy - x - 2y^2 = C$ (Exact)
   (d) $y = (x^2 + \frac{C}{x})^2$ (Bernoulli)
   (e) $y = \frac{x}{C-\ln x}$ (Homogeneous)
   (f) $y = \frac{x}{1+2x\ln x+Cx}$ (Separable)
   (g) $y = \tan\left(\frac{x^3}{3} + x + C\right)$ (Separable)
   (h) $y = \frac{1}{2} - \frac{1}{x} + \frac{C}{x^2}$ (Linear)
   (i) $y = 1 + (1-x)\ln|1-x| + C(1-x)$ (Linear)
   (j) $3x^2y^3 + 2xy^4 = C$ (Exact, homogeneous)
   (k) $y^2 = \frac{x^2}{2\ln|x|+C}$ (Homogeneous, Bernoulli)
   (l) $y = x^{-1}(\ln|x| + C)^{-\frac{1}{3}}$ (Bernoulli)

Exercise 2.1

1. (a) $\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}$
   (b) $\begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$
   (c) $\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix}$
   (d) $\begin{pmatrix} 1 & -4 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$
   (e) $\begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
   (f) $\begin{pmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
   (g) $\begin{pmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
   (h) $\begin{pmatrix} 1 & 2 & 0 & 2 & 3 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
   (i) $\begin{pmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}$

2. (a) $(x_1, x_2, x_3) = (13 + 11\alpha, 2 + 5\alpha, \alpha)$
   (b) Inconsistent
   (c) $(x_1, x_2, x_3) = (7 - 2\alpha, \alpha - 1, \alpha)$
   (d) $(x_1, x_2, x_3, x_4) = (13 + 4\alpha, 6 + \alpha, 5 + 3\alpha, \alpha)$
   (e) $(x_1, x_2, x_3, x_4) = (2\alpha - \beta, \alpha, \beta, 1)$
   (f) $(x_1, x_2, x_3, x_4) = (4 + 2\alpha - 3\beta, \alpha, 3 - 4\beta, \beta)$

Exercise 2.2

1. There are many possible answers. For example, $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

2. There are many possible answers. For example, $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

3. Let $\mathbf{S} = \frac{\mathbf{A}+\mathbf{A}^T}{2}$ and $\mathbf{K} = \frac{\mathbf{A}-\mathbf{A}^T}{2}$. Then $\mathbf{S}^T = \frac{\mathbf{A}^T+(\mathbf{A}^T)^T}{2} = \frac{\mathbf{A}^T+\mathbf{A}}{2} = \mathbf{S}$ and $\mathbf{K}^T = \frac{\mathbf{A}^T-(\mathbf{A}^T)^T}{2} = \frac{\mathbf{A}^T-\mathbf{A}}{2} = -\mathbf{K}$. Now we have $\mathbf{A} = \mathbf{S} + \mathbf{K}$ where $\mathbf{S}$ is symmetric and $\mathbf{K}$ is skew-symmetric.

4. Let $\mathbf{S} = \mathbf{A} - \mathbf{B} = \mathbf{D} - \mathbf{C}$. Observe that $\mathbf{S}$ is both symmetric and skew-symmetric. We must have $\mathbf{S} = \mathbf{0}$. Therefore $\mathbf{A} = \mathbf{B}$ and $\mathbf{C} = \mathbf{D}$.

5.
$$\begin{aligned} & \mathbf{A}^2 - (a+d)\mathbf{A} + (ad-bc)\mathbf{I} \\ = & \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ = & \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} - \begin{pmatrix} a^2+ad & ab+bd \\ ac+cd & ad+d^2 \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} \\ = & \mathbf{0} \end{aligned}$$

6.
$$\begin{aligned} & (\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2 \\ \Leftrightarrow\ & \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2 \\ \Leftrightarrow\ & \mathbf{BA} = \mathbf{AB} \end{aligned}$$

Exercise 2.3

1. (a) $\begin{pmatrix} 5 & -6 \\ -4 & 5 \end{pmatrix}$
   (b) $\frac{1}{2}\begin{pmatrix} 6 & -7 \\ -4 & 5 \end{pmatrix}$
   (c) $\begin{pmatrix} -5 & -2 & 5 \\ 2 & 1 & -2 \\ -4 & -3 & 5 \end{pmatrix}$
   (d) $\begin{pmatrix} 18 & 2 & -7 \\ -3 & 0 & 1 \\ -4 & -1 & 2 \end{pmatrix}$
   (e) $\frac{1}{3}\begin{pmatrix} -3 & 0 & 3 \\ -1 & -3 & -1 \\ -1 & 3 & 2 \end{pmatrix}$
   (f) $\frac{1}{5}\begin{pmatrix} 1 & 2 & -2 \\ -5 & 0 & 5 \\ -3 & -1 & 6 \end{pmatrix}$
   (g) $\frac{1}{5}\begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & -2 & -1 & 2 \\ 0 & 1 & 1 & -1 \\ -3 & 3 & -5 & 1 \end{pmatrix}$

2. (a) $x_1 = 3, x_2 = -1$ (b) $x_1 = 1, x_2 = -11, x_3 = 16$

3. (a) $\begin{pmatrix} 3 & 0 \\ 7 & -1 \\ 1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 12 & -3 \\ -8 & 1 \\ -21 & 9 \end{pmatrix}$

4. We have

$$(\mathbf{I} + \mathbf{B}\mathbf{A}^{-1})\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1} = (\mathbf{A} + \mathbf{B})(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{I}$$

and

$$\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{I} + \mathbf{B}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{A} + \mathbf{B})\mathbf{A}^{-1} = \mathbf{A}\mathbf{I}\mathbf{A}^{-1} = \mathbf{I}.$$

Therefore $(\mathbf{I} + \mathbf{B}\mathbf{A}^{-1})^{-1} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}$.

5. We prove the statement by finding an inverse for $\mathbf{I} - \mathbf{A}$. Now

$$\begin{aligned} & (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1}) \\ = \; & (\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1}) - (\mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \cdots + \mathbf{A}^k) \\ = \; & \mathbf{I} - \mathbf{A}^k \\ = \; & \mathbf{I} \end{aligned}$$

Similarly $(\mathbf{I}+\mathbf{A}+\mathbf{A}^2+\cdots+\mathbf{A}^{k-1})(\mathbf{I}-\mathbf{A}) = \mathbf{I}$. Therefore $(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I}+\mathbf{A}+\mathbf{A}^2+\cdots+\mathbf{A}^{k-1}$.

6. We prove the statement by finding the inverse of $\mathbf{A}^{-1} + \mathbf{B}^{-1}$. Since

$$\begin{aligned} (\mathbf{A}^{-1} + \mathbf{B}^{-1})\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B} &= (\mathbf{I} + \mathbf{B}^{-1}\mathbf{A})(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B} \\ &= \mathbf{B}^{-1}(\mathbf{B} + \mathbf{A})(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B} \\ &= \mathbf{B}^{-1}\mathbf{I}\mathbf{B} \\ &= \mathbf{I} \end{aligned}$$

and

$$\begin{aligned} \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}(\mathbf{A}^{-1} + \mathbf{B}^{-1}) &= \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{B}\mathbf{A}^{-1} + \mathbf{I}) \\ &= \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{B} + \mathbf{A})\mathbf{A}^{-1} \\ &= \mathbf{A}\mathbf{I}\mathbf{A}^{-1} \\ &= \mathbf{I} \end{aligned}$$

we have $(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}$.

7. First we have $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Differentiating both sides with respect to t, we have

$$\begin{aligned} \frac{d}{dt}\left(\mathbf{A}^{-1}\mathbf{A}\right) &= \mathbf{0} \\ \left(\frac{d}{dt}\mathbf{A}^{-1}\right)\mathbf{A} + \mathbf{A}^{-1}\frac{d}{dt}\mathbf{A} &= \mathbf{0} \\ \left(\frac{d}{dt}\mathbf{A}^{-1}\right)\mathbf{A} &= -\mathbf{A}^{-1}\frac{d}{dt}\mathbf{A} \\ \frac{d}{dt}\mathbf{A}^{-1} &= -\mathbf{A}^{-1}\left(\frac{d}{dt}\mathbf{A}\right)\mathbf{A}^{-1} \end{aligned}$$

8. Suppose $\mathbf{x}$ is a vector such that $\mathbf{Ax} = \mathbf{0}$, then $\mathbf{x}^T\mathbf{Ax} = 0$ and $\mathbf{x}^T\mathbf{A}^T\mathbf{x} = (\mathbf{Ax})^T\mathbf{x} = 0$. Now we have

$$\begin{aligned}
(\mathbf{Sx})^T(\mathbf{Sx}) &= \mathbf{x}^T\mathbf{S}^T\mathbf{Sx} \\
&= \mathbf{x}^T\mathbf{S}^2\mathbf{x} \\
&= \mathbf{x}^T(\mathbf{A}+\mathbf{A}^T)\mathbf{x} \\
&= \mathbf{x}^T\mathbf{Ax}+\mathbf{x}^T\mathbf{A}^T\mathbf{x} \\
&= \mathbf{x}^T\mathbf{0}+(\mathbf{Ax})^T\mathbf{x} \\
&= 0.
\end{aligned}$$

Thus $\mathbf{Sx} = \mathbf{0}$. Since $\mathbf{S}$ is non-singular, we have $\mathbf{x} = \mathbf{0}$. Therefore $\mathbf{A}$ is non-singular.

Exercise 2.4

1. (a) -1 (b) -58 (c) 0 (d) 10

2. (a) a (c) $(-1)^n a$ (e) $\frac{1}{a}$
   (b) a^2 (d) $3^n a$ (f) a^{n-1}

3. (a) $\begin{pmatrix} 3 & -5 & -5 \\ -3 & 4 & 5 \\ 2 & -2 & -3 \end{pmatrix}$ (b) $\dfrac{1}{2}\begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\dfrac{1}{8}\begin{pmatrix} 2 & -3 & 3 \\ 2 & 1 & -1 \\ -2 & -1 & 9 \end{pmatrix}$

4. (a) $x_1 = 1, x_2 = 1, x_3 = 2$ (c) $x_1 = -\frac{10}{11}, x_2 = \frac{18}{11}, x_3 = \frac{38}{11}$
   (b) $x_1 = \frac{4}{5}, x_2 = -\frac{3}{2}, x_3 = -\frac{8}{5}$ (d) $x_1 = -\frac{144}{55}, x_2 = -\frac{61}{55}, x_3 = \frac{46}{11}$

5. Let $p(a,b,c)$ be the given determinant. First note $p(a,b,c)$ is a polynomial of degree 3. Secondly, if $a = b$, then the $p(a,b,c) = 0$. It follows that $p(a,b,c)$ has a factor $b-a$. Similarly $p(a,b,c)$ has factors $c-b$ and $c-a$. Thus $p(a,b,c) = k(b-a)(c-b)(c-a)$ where k is a constant. Observe that the coefficient of bc^2 is 1. Therefore $k = 1$ and $p(a,b,c) = (b-a)(c-b)(c-a)$.

6.

$$\begin{aligned}
\frac{d}{dt}\begin{vmatrix} a(t) & b(t) \\ c(t) & d(t) \end{vmatrix} &= \frac{d}{dt}(a(t)d(t)-b(t)c(t)) \\
&= a'(t)d(t)+a(t)d'(t)-b'(t)c(t)-b(t)c'(t) \\
&= a'(t)d(t)-b'(t)c(t)+a(t)d'(t)-b(t)c'(t) \\
&= \begin{vmatrix} a'(t) & b'(t) \\ c(t) & d(t) \end{vmatrix} + \begin{vmatrix} a(t) & b(t) \\ c'(t) & d'(t) \end{vmatrix}
\end{aligned}$$

Exercise 2.5

1. (a) $y = x^2-5$ (b) $y = x^2-x+3$ (c) $y = \frac{1}{2}x^2-\frac{3}{2}x$

2. (a) $x^2+y^2-6x-4y-12 = 0$ (b) $x^2+y^2-6x-8y-75 = 0$ (c) $x^2+y^2+4x+4y-5 = 0$

3. $y = -x^3+3x+5$

4.

(a) $y = 3 + \frac{2}{x}$

(b) $y = 10x + \frac{8}{x} - \frac{16}{x^2}$

(c) $y = \frac{6}{x+2}$

(d) $y = \frac{3x-4}{x-2}$

Exercise 3.2

1. (a) Yes (b) No (c) No (d) Yes (e) Yes (f) No

2. (a) Yes (b) Yes (c) No (d) Yes (e) Yes (f) Yes

3. (a) Yes (b) No (c) No (d) Yes

Exercise 3.3

1. (a) Yes (b) No (c) Yes (d) No (e) No (f) Yes

2. (a) Yes (b) Yes (c) No (d) No

3. Suppose $c_1(\mathbf{v}_1+\mathbf{v}_2)+c_2(\mathbf{v}_2+\mathbf{v}_3)+c_3(\mathbf{v}_1+\mathbf{v}_3) = \mathbf{0}$. Then $(c_1+c_3)\mathbf{v}_1+(c_2+c_1)\mathbf{v}_2+(c_3+c_2)\mathbf{v}_3 = \mathbf{0}$. Since $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent, we have $c_1+c_3 = c_2+c_1 = c_3+c_2 = 0$. This implies $c_1 = ((c_1+c_2)+(c_1+c_3)-(c_2+c_3))/2 = 0$. Similarly, $c_2 = c_3 = 0$. Therefore $\mathbf{v}_1+\mathbf{v}_2, \mathbf{v}_2+\mathbf{v}_3, \mathbf{v}_1+\mathbf{v}_3$ are linearly independent.

4. Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ be a set of k vectors. Suppose one of the vectors is zero. Without loss of generality, we may assume $\mathbf{v} = \mathbf{0}$. Then $1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \cdots + 0 \cdot \mathbf{v}_k = \mathbf{0}$ and not all of the coefficients of the linear combination are zero. Therefore $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly dependent.

5. Let $T = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be a set of linearly independent vectors. Let S be a subset of T. Without loss of generality, we may assume $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_l\}$ for some $l \leq k$. To prove that the set S is linearly independent, suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_l\mathbf{v}_l = \mathbf{0}$. Then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_l\mathbf{v}_l + 0 \cdot \mathbf{v}_{l+1} + \cdots + 0 \cdot \mathbf{v}_k = \mathbf{0}$. Since T is linearly independent, all coefficients $c_1, c_2, \cdots, c_l, 0, \cdots, 0$ in the linearly combination are zero. Thus $c_1 = c_2 = \cdots = c_l = 0$. Therefore S is linearly independent.

6. Suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + c\mathbf{v} = \mathbf{0}$. Then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = -c\mathbf{v}$. Now $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \in W$ which implies $-c\mathbf{v} \in W$. However since $\mathbf{v} \neq W$, we must have $c = 0$. It follows that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$. Hence $c_1 = c_2 = \cdots = c_k = 0$ since $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent. Therefore $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k, \mathbf{v}$ are linearly independent.

Exercise 3.4

1. The answer is not unique.

   (a) $\{(2, -1, 0), (4, 0, 1)\}$ (b) $\{(1, 0, 3), (0, 1, -1)\}$ (c) $\{(1, -3, 0), (0, 0, 1)\}$

2. The answer is not unique.

(a) $\{(11, 7, 1)\}$

(b) $\{(11, -5, 1)\}$

(c) $\{(-11, -3, 1, 0), (-11, -5, 0, 1)\}$

(d) $\{(3, -2, 1, 0), (-4, -3, 0, 1)\}$

(e) $\{(2, -3, 1, 0)\}$

(f) $\{(1, -3, 1, 0), (-2, 1, 0, 1)\}$

Exercise 3.5

1. The answer is not unique.

   (a) Null space: $\{(-11, 4, 1)^T\}$; Row space: $\{(1, 0, 11), (0, 1, -4)\}$;
   Column space: $\{(1, 1, 2)^T, (2, 5, 5)^T\}$

   (b) Null space: $\{(2, -3, 1, 0)^T\}$; Row space: $\{(1, 0, -2, 0), (0, 1, 3, 0), (0, 0, 0, 1)\}$;
   Column space: $\{(1, 3, 2)^T, (1, 1, 5)^T, (1, 4, 12)^T\}$

   (c) Null space: $\{(2, 1, 0, 0, 0)^T, (-7, 0, 2, 5, 1)^T\}$;
   Row space: $\{(1, -2, 0, 0, 7), (0, 0, 1, 0, -2), (0, 0, 0, 1, -5)\}$;
   Column space: $\{(3, 1, 1)^T, (1, 0, 2)^T, (3, 1, 0)^T\}$

   (d) Null space: $\{(3, -2, 1, 0)^T, (-4, -3, 0, 1)^T\}$; Row space: $\{(1, 0, -3, 4), (0, 1, 2, 3)\}$;
   Column space: $\{(1, 1, 1, 2)^T, (1, 4, 3, 5)^T\}$

   (e) Null space: $\{(-1, -2, 1, 0)^T\}$; Row space: $\{(1, 0, 1, 0), (0, 1, 2, 0), (0, 0, 0, 1)\}$;
   Column space: $\{(1, 1, 1, 2)^T, (-2, 4, 3, 2)^T, (-5, 2, 1, -3)^T\}$

   (f) Null space: $\{(-2, -1, 1, 0, 0)^T, (-1, -2, 0, 1, 0)^T\}$;
   Row space: $\{(1, 0, 2, 1, 0), (0, 1, 1, 2, 0), (0, 0, 0, 0, 1)\}$;
   Column space: $\{(1, 2, 2, 3)^T, (1, 3, 3, 1)^T, (1, 2, 3, 4)^T\}$

   (g) Null space: $\{(-2, -1, 1, 0, 0)^T, (-1, -2, 0, 0, 1)^T\}$;
   Row space: $\{(1, 0, 2, 0, 1), (0, 1, 1, 0, 2), (0, 0, 0, 1, 0)\}$;
   column space: $\{(1, -1, 2, -2)^T, (1, 0, 3, 4)^T, (0, 1, 1, 7)^T\}$

2. The answer is not unique.

   (a) $\{(1, 0, 1, 1), (0, 1, -1, 1)\}$

   (b) $\{(1, 0, 2, 2), (0, 1, 0, -1)\}$

   (c) $\{(1, 0, 2, 0), (0, 1, -2, 0), (0, 0, 0, 1)\}$

   (d) $\{(1, -2, 1, 1, 2), (0, 1, 1, 3, 0), (0, 0, 0, 0, 1)\}$

   (e) $\{(1, -3, 4, -2, 5), (0, 0, 1, 3, -2), (0, 0, 0, 0, 1)\}$

3. Denote by n_A, n_B and n_{AB} the nullity of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{AB}$ respectively. By the rank nullity theorem, we have $r_A = n - n_A$, $r_B = k - n_B$ and $r_{AB} = k - n_{AB}$. Now we have (Theorem 3.5.10)

   $$\begin{array}{rcl} n_B \leq & n_{AB} & \leq n_A + n_B \\ k - r_B \leq & k - r_{AB} & \leq (n - r_A) + (k - r_B) \\ r_B \geq & r_{AB} & \geq r_A + r_B - n \end{array}$$

   Moreover, we have

   $$r_{AB} = \text{rank}((\mathbf{AB})^T) = \text{rank}(\mathbf{B}^T\mathbf{A}^T) \leq \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A}) = r_A$$

   Therefore

   $$r_A + r_B - n \leq r_{AB} \leq \min(r_A, r_B)$$

Exercise 3.6

1. The answer is not unique.
   (a) $\{(2,1,0),(-3,0,1)\}$
   (b) $\{(2,1,0,0),(3,0,1,0),(-5,0,0,1)\}$
   (c) $\{(7,-3,1,0),(-19,5,0,1)\}$
   (d) $\{(-12,-3,1,0),(16,7,0,1)\}$
   (e) $\{(-13,4,1,0,0),(4,-3,0,1,0),(-11,4,0,0,1)\}$
   (f) $\{(-5,1,1,0,0),(-12,4,0,1,0),(-19,7,0,0,1)\}$
   (g) $\{(-1,-1,1,0,0),(0,-1,0,-1,1)\}$
   (h) $\{(-2,1,1,0,0),(1,-2,0,1,0)\}$

2. (a)

$$\begin{aligned} |\mathbf{u}+\mathbf{v}|^2+|\mathbf{u}-\mathbf{v}|^2 &= \langle \mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}\rangle + \langle \mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v}\rangle \\ &= \langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle + \langle \mathbf{u},\mathbf{u}\rangle - 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \\ &= 2|\mathbf{u}|^2+2|\mathbf{v}|^2 \end{aligned}$$

   (b)

$$\begin{aligned} |\mathbf{u}+\mathbf{v}|^2-|\mathbf{u}-\mathbf{v}|^2 &= \langle \mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}\rangle + \langle \mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v}\rangle \\ &= (\langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle) - (\langle \mathbf{u},\mathbf{u}\rangle - 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle) \\ &= 4\langle \mathbf{u},\mathbf{v}\rangle \end{aligned}$$

3. Suppose $\mathbf{w}\in W\cap W^{\perp}$. Consider $\langle \mathbf{w},\mathbf{w}\rangle$. Note that the first vector is in W and the second vector is in $W^{\perp}$. We must have $\langle \mathbf{w},\mathbf{w}\rangle=0$. Therefore $\mathbf{w}=\mathbf{0}$ and $W\cap W^{\perp}=\{\mathbf{0}\}$.

4. Suppose $\mathbf{w}\in W$. For any $\mathbf{v}\in W^{\perp}$, we have $\langle \mathbf{w},\mathbf{v}\rangle=0$ since $\mathbf{w}\in W$ and $\mathbf{v}\in W^{\perp}$. Hence $\mathbf{w}\in (W^{\perp})^{\perp}$. Therefore $W\subset (W^{\perp})^{\perp}$.

5. Suppose $c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_k\mathbf{v}_k=\mathbf{0}$. For any $i=1,2,\ldots,k$, we have $c_1\langle \mathbf{v}_i,\mathbf{v}_1\rangle + c_2\langle \mathbf{v}_i,\mathbf{v}_2\rangle+\cdots+c_k\langle \mathbf{v}_i,\mathbf{v}_k\rangle = \langle \mathbf{v}_i,\mathbf{0}\rangle=0$. Since $\langle \mathbf{v}_i,\mathbf{v}_j\rangle=0$ whenever $i\neq j$, we get $c_i\langle \mathbf{v}_i,\mathbf{v}_i\rangle=0$. Since $\mathbf{v}_i$ is a non-zero vector, we obtain $c_i=0$ for $i=1,2,\cdots,k$. Therefore $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k$ are linearly independent.

Exercise 4.1

1. (a) linearly dependent
   (b) linearly independent
   (c) linearly independent
   (d) linearly dependent
   (e) linearly independent
   (f) linearly dependent
   (g) linearly dependent
   (h) linearly dependent

2. (a) $-5e^{-t}$
   (b) 1
   (c) e^{4t}
   (d) t^2e^t
   (e) e^{2t}
   (f) 0

3. $\frac{3}{25}$

4. $3e^{\frac{2}{3}}$

5.

$$\begin{aligned} W[y_1,y_2](t) &= \begin{vmatrix} t^3 & |t^3| \\ 3t^2 & 3t|t| \end{vmatrix} \\ &= 3t^4|t| - 3t^2|t^3| \\ &= 0 \end{aligned}$$

Suppose y_1 and y_2 are solution to a second order linear equation. Then Theorem 4.1.14 asserts that y_1 and y_2 must be linearly dependent since their Wronskian is zero. This contradicts that fact that y_1 and y_2 are linearly independent. Therefore the assumption that y_1 and y_2 are solution to a second order linear equation cannot be true.

6.

$$\begin{aligned} W[fg,fh](t) &= \begin{vmatrix} fg & fh \\ f'g+fg' & f'h+fh' \end{vmatrix} \\ &= fg(f'h+fh') - fh(f'g+fg') \\ &= f^2gh' - f^2hg' \\ &= f^2(gh'-hg') \\ &= f^2\begin{vmatrix} g & h \\ g' & h' \end{vmatrix} \\ &= f^2W[g,h](t) \end{aligned}$$

Exercise 4.2

1. (a) $y = c_1t^2 + c_2t^{-1}$
   (b) $y = c_1t^{-1} + c_2t^{-2}$
   (c) $y = c_1t + c_2t\ln t$
   (d) $y = c_1e^t + c_2te^t$
   (e) $y = c_1\cos(t^2) + c_2\sin(t^2)$
   (f) $y = c_1t^{-1} + c_2t^{-1}\ln t$
   (g) $y = c_1t^2 + c_2t^2\ln t$
   (h) $y = c_1\cos(\ln t) + c_2\sin(\ln t)$

Exercise 4.3

1. (a) $y = c_1e^{-3t} + c_2e^{2t}$
   (b) $y = c_1\cos 3t + c_2\sin 3t$
   (c) $y = c_1e^{2t} + c_2e^t$
   (d) $y = c_1te^{4t} + c_2e^{4t}$
   (e) $y = e^{-2t}(c_1\cos 3t + c_2\sin 3t)$
   (f) $y = e^t(c_1\cos 2t + c_2\sin 2t)$

2. (a) $y = e^t$
   (b) $y = -1 - e^{-3t}$
   (c) $y = 9e^{-2t} - 7e^{-3t}$
   (d) $y = 4e^{-\frac{t}{2}}\cos t + 3e^{-\frac{t}{2}}\sin t$

3. (a) $y = c_1t^{-4} + c_2t^3$
   (b) $y = c_1t^2 + c_2t^2\ln x$

Exercise 4.4

1. (a) $y = c_1e^{3t} + c_2e^{-t} - e^{2t}$
   (b) $y = c_1e^{-t}\cos 2t + c_2e^{-t}\sin 2t + \frac{3}{17}\cos 2t + \frac{12}{17}\sin 2t$
   (c) $y = c_1\cos 3t + c_2\sin 3t + \frac{1}{162}(9t^2 - 6t + 1)e^{3t} + \frac{2}{3}$
   (d) $y = c_1e^t + c_2e^{-2t} - t - \frac{1}{2}$

- (e) $y = c_1e^{-t} + c_2te^{-t} + t^2e^{-t}$
- (f) $y = c_1e^t + c_2te^t + \frac{1}{6}t^3e^t + 4$
- (g) $y = c_1 \cos 2t + c_2 \sin 2t + \frac{1}{4}t^2 - \frac{1}{8} + \frac{3}{5}e^t$
- (h) $y = c_1e^{-t} \cos 2t + c_2e^{-t} \sin 2t + te^{-t} \sin 2t$
- (i) $y = c_1e^{-t} + c_2e^{4t} - e^{2t} + \frac{3}{17} \cos t - \frac{5}{17} \sin t$
- (j) $y = C_1e^{-2t} + C_2te^{-2t} + \frac{1}{2}e^{2t}4t^2e^{-2t}$

2. (a) $y_p = t(A_4t^4 + A_3t^3 + A_2t^2 + A_1t + A_0) + t(B_2t^2 + B_1t + B_0)e^{-3t} + D \cos 3t + E \sin 3t$
   - (b) $y_p = e^t(A_1 \cos 2t + A_2 \sin 2t) + (B_1t + B_0)e^{2t} \cos t + (E_1t + E_0)e^{2t} \sin t$
   - (c) $y_p = A_1t + A_0 + t(B_1t + B_0) \cos t + t(C_1t + C_0) \sin t$
   - (d) $y_p = Ae^{-t} + t(B_2t^2 + B_1t + B_0)e^{-t} \cos t + t(C_2t^2 + C_1t + C_0)e^{-t} \sin t$
   - (e) $y_p = A_1t + A_0 + t^2(B_1t + B_0) + (C_1t + C_0) \cos t + (D_1t + D_0) \sin t$

Exercise 4.5

1. (a) $y = c_1e^{2t} + c_2e^{3t} + e^t$
   - (b) $y = c_1e^{-t} + c_2e^{2t} - \frac{2}{3}te^{-t}$
   - (c) $y = c_1e^{-t} + c_2te^{-t} + 2t^2 + e^{-t}$
   - (d) $y = c_1 \cos t + c_2 \sin t - \cos t \ln(\sec t + \tan t)$
   - (e) $y = c_1 \cos 3t + c_2 \sin 3t + \sin 3t \ln(\sec 3t + \tan 3t) - 1$
   - (f) $y = c_1e^t + c_2te^t - \frac{1}{2}e^t \ln(1 + t^2) + te^t \tan^{-1} t$
   - (g) $y = c_1e^t + c_2e^{2t} - e^{-t} + e^t \ln(1 + e^{-t}) + e^{2t} \ln(1 + e^{-t})$

2. (a) $y = c_1t^{-1} + c_2t^2 + \frac{1}{2} + t^2 \ln t$
   - (b) $y = c_1t + c_2te^t - 2t^2$
   - (c) $y = c_1t + c_2e^t - \frac{1}{2}(2t - 1)e^{-t}$
   - (d) $y = c_1t^2 + c_2t^2 \ln t + \frac{1}{6}t^2(\ln t)^3$

Exercise 4.7

1. (a) $y_p(t) = At + Bt \cos t + Ct \sin t$
   - (b) $y_p(t) = t(A_1t + A_0) + t(B_1t + B_0)e^t \cos t + t(C_1t + C_0)e^t \sin t$
   - (c) $y_p(t) = t^2(A_1t + A_0)e^t$
   - (d) $y_p(t) = Ate^t + t(B_1t + B_0)e^{2t}$
   - (e) $y_p(t) = t^2(A_1t + A_0) \cos t + t^2(B_1t + B_0) \sin t$
   - (f) $y_p(t) = t^2(A_2t^2 + A_1t + A_0)$
   - (g) $y_p(t) = Ate^t + t^3(B_2t^2 + B_1t + B_0)$

2. (a) $-\frac{1}{2}t^2$
   - (b) $y = \frac{1}{6}t^2e^{2t}$
   - (c) $y = \frac{1}{30}e^{4t}$
   - (d) $y = \ln(\sec t) - \sin t \ln(\sec t + \tan t)$

Exercise 5.1

1. (a) $\lambda_1 = -1$, $\{\mathbf{v}_1 = (1, 1)^T\}$; $\lambda_2 = 2$, $\{\mathbf{v}_2 = (2, 1)^T\}$
   - (b) $\lambda_1 = 2$, $\{\mathbf{v}_1 = (1, 1)^T\}$; $\lambda_2 = 4$, $\{\mathbf{v}_2 = (3, 2)^T\}$
   - (c) $\lambda_1 = i$, $\{\mathbf{v}_1 = (-1, 2 + i)^T\}$; $\lambda_2 = -i$, $\{\mathbf{v}_2 = (-1, 2 - i)^T\}$

(d) $\lambda_1 = \lambda_2 = 2$, $\{\mathbf{v}_1 = (1,1)^T\}$

(e) $\lambda_1 = 0$, $\{\mathbf{v}_1 = (1,1,1)^T\}$; $\lambda_2 = 1$, $\{\mathbf{v}_2 = (3,2,1)^T\}$; $\lambda_3 = 2$, $\{\mathbf{v}_3 = (7,3,1)^T\}$

(f) $\lambda_1 = 2$, $\{\mathbf{v}_1 = (1,1,0)^T\}$; $\lambda_2 = \lambda_3 = 1$, $\{\mathbf{v}_2 = (1,0,0)^T, \mathbf{v}_3 = (0,1,-1)^T\}$

(g) $\lambda_1 = \lambda_2 = \lambda_3 = -1$, $\{\mathbf{v}_1 = (1,0,1)^T\}$

(h) $\lambda_1 = 3$, $\{\mathbf{v}_1 = (1,0,0)^T\}$; $\lambda_2 = \lambda_3 = 1$, $\{\mathbf{v}_2 = (1,0,1)^T, \mathbf{v}_3 = (-3,1,0)^T\}$

(i) $\lambda_1 = 1$, $\{\mathbf{v}_1 = (1,-1,8)^T\}$; $\lambda_2 = \lambda_3 = 2$, $\{\mathbf{v}_2 = (0,0,1)^T\}$

(j) $\lambda_1 = \lambda_2 = \lambda_3 = -1$, $\{\mathbf{v}_1 = (1,1,0)^T, \mathbf{v}_2 = (-5,3,8)^T\}$

2. Suppose λ is an eigenvalue of $\mathbf{A}$. Then there exists $\xi \neq \mathbf{0}$ such that $\mathbf{A}\xi = \lambda\xi$.

$$\begin{aligned}
\mathbf{A}^2\xi &= \mathbf{A}\xi \\
\mathbf{A}(\lambda\xi) &= \lambda\xi \\
\lambda\mathbf{A}\xi &= \lambda\xi \\
\lambda^2\xi &= \lambda\xi \\
(\lambda^2 - \lambda)\xi &= \mathbf{0}
\end{aligned}$$

Thus $\lambda^2 - \lambda = 0$ since $\xi \neq \mathbf{0}$. Therefore $\lambda = 0$ or 1.

3. (a) Since $\det(\mathbf{A}^T - \lambda\mathbf{I}) = \det(\mathbf{A} - \lambda\mathbf{I})$, the characteristic equation of $\mathbf{A}$ and $\mathbf{A}^T$ are the same. Therefore $\mathbf{A}^T$ and $\mathbf{A}$ have the same set of eigenvalues.

(b) The matrix $\mathbf{A}$ is non-singular if and only if $\det(\mathbf{A}) = 0$ if and only if $\lambda = 0$ is a root of the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ if and only if $\lambda = 0$ is an eigenvalue of $\mathbf{A}$.

(c) Suppose λ is an eigenvalue of $\mathbf{A}$. Then there exists $\xi \neq \mathbf{0}$ such that $\mathbf{A}\xi = \lambda\xi$. Now

$$\mathbf{A}^k\xi = \mathbf{A}^{k-1}\mathbf{A}\xi = \mathbf{A}^{k-1}(\lambda\xi) = \lambda\mathbf{A}^{k-1}\xi = \cdots = \lambda^k\xi.$$

Therefore λ^k is an eigenvalue of $\mathbf{A}^k$.

(d) Suppose λ is an eigenvalue of a non-singular matrix $\mathbf{A}$. Then there exists $\xi \neq \mathbf{0}$ such that

$$\begin{aligned}
\mathbf{A}\xi &= \lambda\xi \\
\xi &= \mathbf{A}^{-1}(\lambda\xi) \\
\xi &= \lambda\mathbf{A}^{-1}\xi \\
\lambda^{-1}\xi &= \mathbf{A}^{-1}\xi
\end{aligned}$$

Therefore λ^{-1} is an eigenvalue of ξ.

4. Let

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ 0 & & & a_{nn} \end{pmatrix},$$

be an upper-triangular matrix. Then λ is an eigenvalue of $\mathbf{A}$ if and only if

$$\begin{aligned}\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ & a_{22}-\lambda & a_{23} & \cdots & a_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & a_{n-1n} \\ & \text{\huge 0} & & & a_{nn}-\lambda \end{vmatrix} &= 0 \\ (\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn}) &= 0 \\ \lambda &= a_{11}, a_{22}, \cdots, a_{nn}\end{aligned}$$

Exercise 5.2

1. (a) $\mathbf{P}=\begin{pmatrix} 3 & -1 \\ 4 & 1 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$
   (b) $\mathbf{P}=\begin{pmatrix} 1 & 1 \\ 1-i & 1+i \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 1+2i & 0 \\ 0 & 1-2i \end{pmatrix}$
   (c) $\mathbf{P}=\begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$
   (d) $\mathbf{P}=\begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$
   (e) $\mathbf{P}=\begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$
   (f) $\mathbf{P}=\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$
   (g) $\mathbf{P}=\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}$, $\mathbf{D}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

2. (a) The characteristic equation of the matrix is $(\lambda-2)^2=0$. There is only one eigenvalue $\lambda=2$. For eigenvalue $\lambda=2$, the eigenspace is $\text{span}((1,1)^T)$ which is of dimension 1. Therefore the matrix is not diagonalizable.
   (b) The characteristic equation of the matrix is $(\lambda-2)(\lambda-1)^2=0$. The algebraic multiplicity of eigenvalue $\lambda=1$ is 2 but the associated eigenspace is spanned by one vector $(1,2,-1)^T$. Therefore the matrix is not diagonalizable.
   (c) The characteristic equation of the matrix is $(\lambda-2)^2(\lambda-1)=0$. The algebraic multiplicity of eigenvalue $\lambda=2$ is 2 but the associated eigenspace is spanned by one vector $(1,1,-1)^T$. Therefore the matrix is not diagonalizable.

3. Suppose $\mathbf{A}$ is similar to $\mathbf{A}$. Then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B}=\mathbf{P}^{-1}\mathbf{AP}$. Then $\mathbf{B}^{-1}=(\mathbf{P}^{-1}\mathbf{AP})^{-1}=\mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{P}$. Therefore $\mathbf{A}^{-1}$ is similar to $\mathbf{B}^{-1}$.

4. The answer is negative. Consider $\mathbf{A}=\mathbf{B}=\mathbf{C}=\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{D}=\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Then $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ are all similar. But $\mathbf{AC}=\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$ is not similar to $\mathbf{BD}=\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

5. If $\mathbf{A}$ is similar to $\mathbf{B}$, then $\mathbf{A}-\lambda\mathbf{I}$ is similar to $\mathbf{B}-\lambda\mathbf{I}$ for any λ. Thus $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic equation. Therefore $\mathbf{A}$ and $\mathbf{B}$ have the same set of eigenvalues.

6. Let $\mathbf{A}=[a_{ij}]$ and $\mathbf{B}=[b_{ij}]$

$$\begin{aligned}\mathrm{tr}(\mathbf{AB}) &= \sum_{k=0}^{n}[\mathbf{AB}]_{kk}\\ &= \sum_{k=0}^{n}\sum_{l=0}^{n}a_{kl}b_{lk}\\ &= \sum_{l=0}^{n}\sum_{k=0}^{n}b_{lk}a_{kl}\\ &= \sum_{l=0}^{n}[\mathbf{BA}]_{ll}\\ &= \mathrm{tr}(\mathbf{BA})\end{aligned}$$

7. The characteristic equation of the matrix is $\lambda^2-(a+d)\lambda+ad-bc=0$. The discriminant of the equation is

$$(a+d)^2-4(ad-bc)=a^2-2ad+d^2+4bc=(a-d)^2+4bc.$$

Hence if the discriminant $(a-d)^2+4bc\neq 0$, then the equation has two distinct roots. This implies that the matrix has two distinct eigenvalues. Their associated eigenvectors are linearly independent. Therefore the matrix is diagonalizable.

8. We have $\mathbf{P}^{-1}\mathbf{AP}=\mathbf{D}_1$ and $\mathbf{P}^{-1}\mathbf{BP}=\mathbf{D}_2$ are diagonal matrices. Thus

$$\begin{aligned}\mathbf{AB} &= (\mathbf{PD}_1\mathbf{P}^{-1})(\mathbf{PD}_2\mathbf{P}^{-1})\\ &= \mathbf{PD}_1\mathbf{D}_2\mathbf{P}^{-1}\\ &= \mathbf{PD}_2\mathbf{D}_1\mathbf{P}^{-1}\\ &= (\mathbf{PD}_2\mathbf{P}^{-1})(\mathbf{PD}_1\mathbf{P}^{-1})\\ &= \mathbf{BA}\end{aligned}$$

9. Suppose $\mathbf{A}$ is similar to a diagonal matrix $\mathbf{D}$. Then $\mathbf{D}^k$ is similar to $\mathbf{A}^k=\mathbf{0}$. This implies that $\mathbf{D}=\mathbf{0}$. It follows that $\mathbf{A}=\mathbf{0}$ since the only matrix similar to the zero matrix is the zero matrix.

10. We have $\mathbf{A}^{-1}(\mathbf{AB})\mathbf{A}=\mathbf{BA}$. Thus $\mathbf{AB}$ is similar to $\mathbf{BA}$.

11. Let $\mathbf{A}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$ and $\mathbf{B}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$. Then $\mathbf{AB}=\mathbf{A}\neq\mathbf{0}$ but $\mathbf{BA}=\mathbf{0}$. So $\mathbf{AB}$ and $\mathbf{BA}$ are not be similar.

12. We claim that the matrices

$$\mathbf{P}=\begin{pmatrix}\mathbf{0}&\mathbf{0}\\ \mathbf{A}&\mathbf{AB}\end{pmatrix}\text{ and }\mathbf{Q}=\begin{pmatrix}\mathbf{BA}&\mathbf{0}\\ \mathbf{A}&\mathbf{0}\end{pmatrix}$$

are similar. Then by Theorem 5.2, the matrices $\mathbf{P}$ and $\mathbf{Q}$ have the same characteristic polynomial which implies that $\mathbf{AB}$ and $\mathbf{BA}$ have the same characteristic polynomial. To prove the claim, we have

$$\begin{aligned}\begin{pmatrix}\mathbf{I} & \mathbf{B}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{BA} & \mathbf{0}\\ \mathbf{A} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{B}\\ \mathbf{0} & \mathbf{I}\end{pmatrix} &= \begin{pmatrix}\mathbf{I} & -\mathbf{B}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{BA} & \mathbf{0}\\ \mathbf{A} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{B}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}\\ &= \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{A} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{B}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}\\ &= \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{A} & \mathbf{AB}\end{pmatrix}.\end{aligned}$$

Caution! In general, **AB** and **BA** may not be similar.

Exercise 5.3

1. (a) $\begin{pmatrix}65 & -66\\ 33 & -34\end{pmatrix}$ (b) $\begin{pmatrix}96 & -96\\ 64 & -64\end{pmatrix}$ (c) $\begin{pmatrix}94 & -93\\ 62 & -61\end{pmatrix}$ (d) $\begin{pmatrix}16 & -80\\ 16 & -16\end{pmatrix}$ (e) $\begin{pmatrix}16 & 32 & -16\\ 32 & 64 & -32\\ 48 & 96 & -48\end{pmatrix}$ (f) $\begin{pmatrix}1 & -62 & 31\\ 0 & 1 & 0\\ 0 & -62 & 32\end{pmatrix}$ (g) $\begin{pmatrix}94 & -93 & 31\\ 62 & -61 & 31\\ 0 & 0 & 32\end{pmatrix}$

2. (a) Observe that $\mathbf{A}^T\xi = \xi$ where $\xi = (1, \cdots, 1)^T$, we have $\mathbf{A}^T$ has an eigenvalue $\lambda = 1$. Therefore $\mathbf{A}$ has an eigenvalue $\lambda = 1$.

   (b) Write $\mathbf{A} = [a_{ij}]$, $1 \le i, j \le n$. Let $\xi = (x_1, x_2, \cdots, x_n)^T$ be an eigenvector of $\mathbf{A}^T$ associated with $\lambda = 1$ and k be such that x_k is maximum among $x_1, x_2, \cdots, x_n$. In other words, $x_k \ge x_l$ for any $l = 1, 2, \cdots, n$. Now consider the k-th row of $\mathbf{A}^T\xi = \xi$ and observe that the sum of the entries of each row of $\mathbf{A}^T$ is 1, we have

$$\begin{aligned} x_k &= a_{1k}x_1 + a_{2k}x_2 + \cdots + a_{nk}x_n\\ &\le a_{1k}x_k + a_{2k}x_k + \cdots + a_{nk}x_k\\ &= (a_{1k} + a_{2k} + \cdots + a_{nk})x_k\\ &= x_k.\end{aligned}$$

   Thus the equality holds on above. It follows that for any $l = 1, 2, \cdots, n$, we have $a_{lk}x_l = a_{lk}x_k$ which implies that $x_l = x_k$ since $a_{lk} \ne 0$. So ξ is a multiple of $(1, \cdots, 1)^T$ and thus the eigenspace of $\mathbf{A}^T$ associated with $\lambda = 1$ is of dimension 1. Therefore the eiqenspace of $\mathbf{A}$ associated with $\lambda = 1$ is of dimension 1.

Exercise 5.4

1. (a) Minimal polynomial: $(x-1)(x-2)$, $\mathbf{A}^4 = 15\mathbf{A} - 14\mathbf{I}$, $\mathbf{A}^{-1} = -\frac{1}{2}\mathbf{A} + \frac{3}{2}\mathbf{I}$
   (b) Minimal polynomial: $(x-1)^2$, $\mathbf{A}^4 = 4\mathbf{A} - 3\mathbf{I}$, $\mathbf{A}^{-1} = -\mathbf{A} + 2\mathbf{I}$
   (c) Minimal polynomial: $x^2 - 2x + 5$, $\mathbf{A}^4 = -12\mathbf{A} + 5\mathbf{I}$, $\mathbf{A}^{-1} = -\frac{1}{5}\mathbf{A} + \frac{2}{5}\mathbf{I}$
   (d) Minimal polynomial: $(x-2)(x-1)^2$, $\mathbf{A}^4 = 11\mathbf{A}^2 - 18\mathbf{A} + 8\mathbf{I}$, $\mathbf{A}^{-1} = \frac{1}{2}\mathbf{A}^2 - 2\mathbf{A} + \frac{5}{2}\mathbf{I}$
   (e) Minimal polynomial: $(x-4)(x-2)$, $\mathbf{A}^4 = 120\mathbf{A} - 224\mathbf{I}$, $\mathbf{A}^{-1} = -\frac{1}{8}\mathbf{A} + \frac{3}{4}\mathbf{I}$
   (f) Minimal polynomial: $(x-1)^2$, $\mathbf{A}^4 = 4\mathbf{A} - 3\mathbf{I}$, $\mathbf{A}^{-1} = -\mathbf{A} + 2\mathbf{I}$
   (g) Minimal polynomial: $(x+1)(x-1)$, $\mathbf{A}^4 = \mathbf{I}$, $\mathbf{A}^{-1} = \mathbf{A}$

2. Suppose $\mathbf{A}$ and $\mathbf{B}$ are similar matrices. Then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. Now for any polynomial function $p(x)$, we have $p(\mathbf{B}) = p(\mathbf{P}^{-1}\mathbf{AP}) = \mathbf{P}^{-1}p(\mathbf{A})\mathbf{P}$. It follows that $p(\mathbf{B}) = \mathbf{0}$ if and only if $p(\mathbf{A}) = \mathbf{0}$. Therefore $\mathbf{A}$ and $\mathbf{B}$ have the same minimal polynomial.

3. The minimal polynomial $m(x)$ of $\mathbf{A}$ divides $x^k - 1 = 0$ which has no repeated factor. It follows that $m(x)$ is a product of distinct linear factors. Therefore $\mathbf{A}$ is diagonalizable by Theorem 5.4.5.

4. Since $\mathbf{A}$ is diagonalizable, by Theorem 5.4.5, the minimal polynomial of $\mathbf{A}^2$ is

$$(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_k)$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of $\mathbf{A}^2$. In particular,

$$(\mathbf{A}^2 - \lambda_1\mathbf{I})(\mathbf{A}^2 - \lambda_2\mathbf{I})\cdots(\mathbf{A}^2 - \lambda_k\mathbf{I}) = \mathbf{0}.$$

Thus

$$(\mathbf{A} - \sqrt{\lambda_1}\mathbf{I})(\mathbf{A} + \sqrt{\lambda_1}\mathbf{I})(\mathbf{A} - \sqrt{\lambda_2}\mathbf{I})(\mathbf{A} + \sqrt{\lambda_2}\mathbf{I})\cdots(\mathbf{A} - \sqrt{\lambda_k}\mathbf{I})(\mathbf{A} + \sqrt{\lambda_k}\mathbf{I}) = \mathbf{0}.$$

It follows that the minimal polynomial of $\mathbf{A}$ divides

$$p(x) = (x - \sqrt{\lambda_1})(x + \sqrt{\lambda_1})(x - \sqrt{\lambda_2})(x + \sqrt{\lambda_2})\cdots(x - \sqrt{\lambda_k})(x + \sqrt{\lambda_k}).$$

Since $\mathbf{A}$ is non-singular, the values $\lambda_1, \lambda_2, \cdots, \lambda_k$ are all non-zero and thus $p(x)$ has no repeated factor. Therefore the minimal polynomial of $\mathbf{A}$ is a product of distinct linear factor and $\mathbf{A}$ is diagonalizable by Theorem 5.4.5.

Exercise 6.2

1. (a) $\begin{cases} x_1 = c_1e^{-t} + c_2e^{3t} \\ x_2 = -c_1e^{-t} + c_2e^{3t} \end{cases}$

   (b) $\begin{cases} x_1 = c_1e^{-t} + 3c_2e^{4t} \\ x_2 = -c_1e^{-t} + 2c_2e^{4t} \end{cases}$

   (c) $\begin{cases} x_1 = 5c_1\cos 2t + 5c_2\sin 2t \\ x_2 = (c_1 - 2c_2)\cos 2t + (2c_1 + c_2)\sin 2t \end{cases}$

   (d) $\begin{cases} x_1 = 3e^{2t}(c_1\cos 3t - c_2\sin 3t) \\ x_2 = e^{2t}((c_1 + c_2)\cos 3t + (c_1 - c_2)\sin 3t) \end{cases}$

   (e) $\begin{cases} x_1 = c_1e^{9t} + c_2e^{6t} + c_3 \\ x_2 = c_1e^{9t} - 2c_2e^{6t} \\ x_3 = c_1e^{9t} + c_2e^{6t} - c_3 \end{cases}$

   (f) $\begin{cases} x_1 = c_1e^{6t} + c_2e^{3t} + c_3e^{3t} \\ x_2 = c_1e^{6t} - 2c_2e^{3t} \\ x_3 = c_1e^{6t} + c_2e^{3t} - c_3e^{3t} \end{cases}$

   (g) $\begin{cases} x_1 = c_1e^{t} + c_2(2\cos 2t - \sin 2t) + c_3(\cos 2t + 2\sin 2t) \\ x_2 = -c_1e^{t} - c_2(3\cos 2t + \sin 2t) + c_3(\cos 2t - 3\sin 2t) \\ x_3 = c_2(3\cos 2t + \sin 2t) + c_3(3\cos 2t - \sin 2t) \end{cases}$

(h) $$\begin{cases} x_1 = c_1e^{5t} + (c_2 + 2c_3)e^t \\ x_2 = c_1e^{5t} - c_2e^t \\ x_3 = c_1e^{5t} - c_3e^t \end{cases}$$

2. (a) $$\begin{cases} x_1 = \frac{1}{7}(-e^{-t} + 8e^{6t}) \\ x_2 = \frac{1}{7}(e^{-t} + 6e^{6t}) \end{cases}$$

(b) $$\begin{cases} x_1 = -5e^{3t} + 6e^{4t} \\ x_2 = 6e^{3t} - 6e^{4t} \end{cases}$$

(c) $$\begin{cases} x_1 = -4e^t \sin 2t \\ x_2 = 4e^t \cos 2t \end{cases}$$

(d) $$\begin{cases} x_1 = 4e^{3t} - e^{-t}(4\cos t - \sin t) \\ x_2 = 9e^{3t} - e^{-t}(9\cos t + 2\sin t) \\ x_3 = 17e^{-t}\cos t \end{cases}$$

3. (a) $$\mathbf{x} = c_1e^{3t}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2e^{-2t}\begin{pmatrix} -2 \\ 3 \end{pmatrix}$$

(b) $$\mathbf{x} = c_1\begin{pmatrix} \cos 2t \\ \cos 2t + 2\sin 2t \end{pmatrix} + c_2\begin{pmatrix} \sin 2t \\ -2\cos 2t + \sin 2t \end{pmatrix}$$

(c) $$\mathbf{x} = c_1e^{2t}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2e^{-t}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_3e^{-2t}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

(d) $$\mathbf{x} = c_1e^{3t}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2e^{3t}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_3e^{2t}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Exercise 6.3

1. (a) $$\mathbf{x} = e^{-t}\left(c_1\begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2\begin{pmatrix} 1+2t \\ -2t \end{pmatrix}\right)$$

(b) $$\mathbf{x} = e^{-3t}\left(c_1\begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2\begin{pmatrix} 1+t \\ -t \end{pmatrix}\right)$$

(c) $$\mathbf{x} = e^{2t}\left(c_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2\begin{pmatrix} 1+t \\ t \end{pmatrix}\right)$$

(d) $$\mathbf{x} = e^{-t}\left(c_1\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} -2 \\ -1+t \\ 1 \end{pmatrix} + c_3\begin{pmatrix} 1-2t \\ -t+\frac{1}{2}t^2 \\ t \end{pmatrix}\right)$$

(e) $$\mathbf{x} = e^{-t}\left(c_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} t \\ 2 \\ 1 \end{pmatrix} + c_3\begin{pmatrix} \frac{1}{2}t^2 \\ 1+2t \\ t \end{pmatrix}\right)$$

(f) $$\mathbf{x} = e^{t}\left(c_1\begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} + c_3\begin{pmatrix} 1 \\ -2t \\ 2t \end{pmatrix}\right)$$

(g) $$\mathbf{x} = e^{2t}\left(c_1\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + c_3\begin{pmatrix} 1+t \\ t \\ 0 \end{pmatrix}\right)$$

(h) $\mathbf{x} = e^{2t}\left(c_1\begin{pmatrix}1\\0\\1\end{pmatrix} + c_2\begin{pmatrix}1-t\\-1\\1-t\end{pmatrix} + c_3\begin{pmatrix}1+t-\frac{1}{2}t^2\\-t\\t-\frac{1}{2}t^2\end{pmatrix}\right)$

Exercise 6.4

1. (a) $\frac{1}{7}\begin{pmatrix}3e^{5t}+4e^{-2t} & 3e^{5t}-3e^{-2t}\\4e^{5t}-4e^{-2t} & 4e^{5t}+3e^{-2t}\end{pmatrix}$

   (b) $\begin{pmatrix}2e^{3t}-e^{t} & -2e^{3t}+2e^{t}\\e^{3t}-e^{t} & -e^{3t}+2e^{t}\end{pmatrix}$

   (c) $\begin{pmatrix}-2+3e^{2t} & 3-3e^{2t}\\-2+2e^{2t} & 3-2e^{2t}\end{pmatrix}$

   (d) $\begin{pmatrix}\cos 2t & \sin 2t\\-\sin 2t & \cos 2t\end{pmatrix}$

   (e) $\begin{pmatrix}1 & 3t\\0 & 1\end{pmatrix}$

   (f) $\begin{pmatrix}-e^{2t}+2e^{3t} & 0 & -e^{2t}+e^{3t}\\2e^{2t}-2e^{3t} & e^{t} & -e^{t}+2e^{2t}-e^{3t}\\2e^{2t}-2e^{3t} & 0 & 2e^{2t}-e^{3t}\end{pmatrix}$

   (g) $\begin{pmatrix}1 & t & \frac{t^2}{2}\\0 & 1 & t\\0 & 0 & 1\end{pmatrix}$

   (h) $\begin{pmatrix}1 & -t & t-6t^2\\0 & 1 & 3t\\0 & 0 & 1\end{pmatrix}$

2. (a) $\begin{pmatrix}-5e^{t}+6e^{-3t}\\e^{t}-6e^{-3t}\end{pmatrix}$

   (b) $\begin{pmatrix}2e^{4t}+2e^{-t}\\3e^{4t}-2e^{-t}\end{pmatrix}$

   (c) $\begin{pmatrix}e^{t}+2e^{-t}\\e^{t}-e^{-t}\\-2e^{t}+e^{-t}\end{pmatrix}$

   (d) $\begin{pmatrix}1+t+2t^2\\1+4t\\-2\end{pmatrix}$

3. (a) By Cayley-Hamilton theorem, we have $\mathbf{A}^2 - 2\lambda\mathbf{A} + (\lambda^2+\mu^2)\mathbf{I} = \mathbf{0}$. Thus

$$\begin{aligned}\mu^2\mathbf{J}^2 &= (\mathbf{A}-\lambda\mathbf{I})^2\\ &= \mathbf{A}^2 - 2\lambda\mathbf{A} + \lambda^2\mathbf{I}\\ &= -\mu^2\mathbf{I}\end{aligned}$$

   Therefore $\mathbf{J}^2 = -\mathbf{I}$.

   (b) Now $\mathbf{A} = \lambda\mathbf{I} + \mu\mathbf{J}$. Therefore

$$\begin{aligned}&\exp(\mathbf{A}t)\\ =\ &\exp((\lambda\mathbf{I}+\mu\mathbf{J})t)\\ =\ &\exp(\lambda\mathbf{I}t)\exp(\mu\mathbf{J}t)\\ =\ &e^{\lambda t}\mathbf{I}(\mathbf{I}+\mu\mathbf{J}t+\frac{1}{2!}\mu^2\mathbf{J}^2t^2+\frac{1}{3!}\mu^3\mathbf{J}^3t^3+\frac{1}{4!}\mu^4\mathbf{J}^4t^4+\frac{1}{5!}\mu^5\mathbf{J}^5t^5+\cdots)\\ =\ &e^{\lambda t}(\mathbf{I}+\mu\mathbf{J}t-\frac{1}{2!}\mu^2\mathbf{I}t^2-\frac{1}{3!}\mu^3\mathbf{J}t^3+\frac{1}{4!}\mu^4\mathbf{I}t^4+\frac{1}{5!}\mu^5\mathbf{J}t^5+\cdots)\\ =\ &e^{\lambda t}(\mathbf{I}(1-\frac{1}{2!}\mu^2t^2+\frac{1}{4!}\mu^4t^2+\cdots)+\mathbf{J}(\mu t-\frac{1}{3!}\mu^3t^3+\frac{1}{5!}\mu^5t^5+\cdots))\\ =\ &e^{\lambda t}(\mathbf{I}\cos\mu t+\mathbf{J}\sin\mu t)\end{aligned}$$

   (c) (i) $\begin{pmatrix}\cos t+2\sin t & -5\sin t\\\sin t & \cos t-2\sin t\end{pmatrix}$ (ii) $e^{t}\begin{pmatrix}\cos 2t+\sin 2t & -\sin 2t\\2\sin 2t & \cos 2t-\sin 2t\end{pmatrix}$

Exercise 6.5

1. (a) $\mathbf{J} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{3t}\begin{pmatrix} 1+t & -t \\ t & 1-t \end{pmatrix}$

   (b) $\mathbf{J} = \begin{pmatrix} -3 & 1 \\ 0 & -3 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{-3t}\begin{pmatrix} 1+4t & -4t \\ 4t & 1-4t \end{pmatrix}$

   (c) $\mathbf{J} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{3t}\begin{pmatrix} 1+2t & -t & t \\ t+t^2 & 1-\frac{t^2}{2} & \frac{t^2}{2} \\ -3t+t^2 & 2t-\frac{t^2}{2} & 1-2t+\frac{t^2}{2} \end{pmatrix}$

   (d) $\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{t}\begin{pmatrix} 1-3t & -9t & 0 \\ t & 1+3t & 0 \\ t & 3t & 1 \end{pmatrix}$

   (e) $\mathbf{J} = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{-2t}\begin{pmatrix} 1+t+\frac{t^2}{2} & t+t^2 & t+\frac{3t^2}{2} \\ t-t^2 & 1+4t-2t^2 & 7t-3t^2 \\ -t+\frac{t^2}{2} & -3t+t^2 & 1-5t+\frac{3t^2}{2} \end{pmatrix}$

   (f) $\mathbf{J} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{2t}\begin{pmatrix} 1 & \frac{t^2}{2} & -t+t^2 \\ 0 & 1+2t & 4t \\ 0 & -t & 1-2t \end{pmatrix}$

   (g) $\mathbf{J} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$, $\exp(\mathbf{A}t) = e^{2t}\begin{pmatrix} 1-3t & 3t & -9t \\ 0 & 1 & 0 \\ t & -t & 1+3t \end{pmatrix}$

   (h) $\mathbf{J} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$,

   $$\exp(\mathbf{A}t) = \begin{pmatrix} (3+2t)e^{2t}-2e^{-t} & (4+3t)e^{2t}-4e^{-t} & (6+4t)e^{2t}-6e^{-t} \\ -4te^{2t} & (1-6t)e^{2t} & -8te^{2t} \\ (-1+2t)e^{2t}+e^{-t} & (-2+3t)e^{2t}+2e^{-t} & (-2+4t)e^{2t}+3e^{-t} \end{pmatrix}$$

Exercise 6.6

1. The answers are not unique.

   (a) $\begin{pmatrix} e^{-t} & 2e^{2t} \\ 2e^{-t} & e^{2t} \end{pmatrix}$

   (b) $\begin{pmatrix} e^{-3t} & e^{2t} \\ -4e^{-3t} & e^{2t} \end{pmatrix}$

   (c) $\begin{pmatrix} \cos t + 2\sin t & -5\sin t \\ \sin t & \cos t - 2\sin t \end{pmatrix}$

   (d) $\begin{pmatrix} 4e^{2t} & e^{-t} \\ -e^{2t} & -e^{-t} \end{pmatrix}$

   (e) $e^{-t}\begin{pmatrix} 2\cos 2t & -2\sin 2t \\ \sin 2t & \cos 2t \end{pmatrix}$

   (f) $e^{-2t}\begin{pmatrix} 3 & 3t+1 \\ 3 & 3t \end{pmatrix}$

   (g) $\begin{pmatrix} -4e^{-2t} & 3e^{-t} & 0 \\ 5e^{-2t} & -4e^{-t} & -e^{2t} \\ 7e^{-2t} & -2e^{-t} & e^{2t} \end{pmatrix}$

   (h) $\begin{pmatrix} e^{-2t} & e^{t} & e^{3t} \\ -e^{-2t} & -4e^{t} & 2e^{3t} \\ -e^{-2t} & -e^{t} & e^{3t} \end{pmatrix}$

   (i) $\begin{pmatrix} 0 & e^{3t} & 0 \\ 2e^{5t} & 0 & e^{3t} \\ e^{5t} & -e^{3t} & e^{3t} \end{pmatrix}$

   (j) $e^{2t}\begin{pmatrix} 1 & 1+t & t+\frac{1}{2}t^2 \\ 2 & 2t & 1+t^2 \\ -1 & -t & -\frac{1}{2}t^2 \end{pmatrix}$

   (k) $\begin{pmatrix} 0 & e^{2t} & te^{2t} \\ e^{t} & e^{2t} & te^{2t} \\ e^{t} & 0 & e^{2t} \end{pmatrix}$

   (l) $e^{-2t}\begin{pmatrix} 1 & -2+t & -2t+\frac{t^2}{2} \\ 0 & 1 & 1+t \\ 0 & 1 & t \end{pmatrix}$

2. (a) $\begin{pmatrix} 4e^{2t}-2e^{-t} & -4e^{2t}+4e^{-t} \\ -e^{2t}+2e^{-t} & e^{2t}-4e^{-t} \end{pmatrix}$

   (b) $e^{5t}\begin{pmatrix} -t & 1+5t \\ -1+2t & 3-10t \end{pmatrix}$

   (c) $\begin{pmatrix} 2e^{3t} & 0 & -e^{3t} \\ 2e^{3t}-2e^{5t} & 5e^{3t}-8e^{5t} & -3e^{3t}+4e^{5t} \\ -5e^{5t} & 5e^{3t}-4e^{5t} & -2e^{3t}+2e^{5t} \end{pmatrix}$

   (d) $e^{2t}\begin{pmatrix} 1-2t & 1-2t & 2t \\ -1+3t-t^2 & -2+3t-t^2 & -2t+t^2 \\ -5t+t^2 & 1-5t+t^2 & 2+4t-t^2 \end{pmatrix}$

3. (a) We have $\mathbf{Q\Psi}$ is non-singular since both $\mathbf{Q}$ and $\mathbf{\Psi}$ are non-singular. Now $\mathbf{Q\Psi}$ is a fundamental matrix for the system if and only if

$$\begin{aligned} \frac{d\mathbf{Q\Psi}}{dt} &= \mathbf{AQ\Psi} \\ \Leftrightarrow \quad \mathbf{Q}\frac{d\mathbf{\Psi}}{dt} &= \mathbf{AQ\Psi} \\ \Leftrightarrow \quad \mathbf{QA\Psi} &= \mathbf{AQ\Psi} \\ \Leftrightarrow \quad \mathbf{QA} &= \mathbf{AQ}. \end{aligned}$$

   (b) By differentiating $\mathbf{\Psi}^{-1}\mathbf{\Psi}=\mathbf{I}$, we have

$$\begin{aligned} \frac{d}{dt}(\mathbf{\Psi}^{-1}\mathbf{\Psi}) &= \mathbf{0} \\ \frac{d\mathbf{\Psi}^{-1}}{dt}\mathbf{\Psi}+\mathbf{\Psi}^{-1}\frac{d\mathbf{\Psi}}{dt} &= \mathbf{0} \\ \frac{d\mathbf{\Psi}^{-1}}{dt}\mathbf{\Psi} &= -\mathbf{\Psi}^{-1}\frac{d\mathbf{\Psi}}{dt} \\ \frac{d\mathbf{\Psi}^{-1}}{dt} &= -\mathbf{\Psi}^{-1}\frac{d\mathbf{\Psi}}{dt}\mathbf{\Psi}^{-1}. \end{aligned}$$

   Now

$$\begin{aligned} \frac{d}{dt}(\mathbf{\Psi}^T)^{-1} &= (\frac{d\mathbf{\Psi}^{-1}}{dt})^T \\ &= -(\mathbf{\Psi}^{-1}\frac{d\mathbf{\Psi}}{dt}\mathbf{\Psi}^{-1})^T \\ &= -(\mathbf{\Psi}^{-1}\mathbf{A\Psi\Psi}^{-1})^T \\ &= -(\mathbf{\Psi}^{-1}\mathbf{A})^T \\ &= -\mathbf{A}^T(\mathbf{\Psi}^{-1})^T \\ &= -\mathbf{A}^T(\mathbf{\Psi}^T)^{-1} \end{aligned}$$

   and $(\mathbf{\Psi}^T)^{-1}$ is non-singular. Therefore $(\mathbf{\Psi}^T)^{-1}$ is a fundamental matrix for the system $\mathbf{x}'=-\mathbf{A}^T\mathbf{x}$.

4. Write

$$\mathbf{\Psi}_1 = \left[\begin{array}{cccc} \mathbf{x}^{(1)} & \mathbf{x}^{(2)} & \cdots & \mathbf{x}^{(n)} \end{array}\right] \text{ and } \mathbf{\Psi}_2 = \left[\begin{array}{cccc} \mathbf{y}^{(1)} & \mathbf{y}^{(2)} & \cdots & \mathbf{y}^{(n)} \end{array}\right],$$

then $\{\mathbf{x}^{(1)},\mathbf{x}^{(2)},\cdots,\mathbf{x}^{(n)}\}$ and $\{\mathbf{y}^{(1)},\mathbf{y}^{(2)},\cdots,\mathbf{y}^{(n)}\}$ constitute two fundamental sets of solutions to the system. In particular for any $i=1,2,\cdots,n$,

$$\mathbf{y}^{(i)} = p_{1i}\mathbf{x}^{(1)}+p_{2i}\mathbf{x}^{(2)}+\cdots+p_{ni}\mathbf{x}^{(n)}, \text{ for some constants } p_{1i},p_{2i},\cdots,p_{ni}.$$

Now let

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix},$$

we have $\mathbf{\Psi}_2 = \mathbf{\Psi}_1\mathbf{P}$. The matrix $\mathbf{P}$ must be non-singular, otherwise $\mathbf{\Psi}_2$ cannot be non-singular.

Alternative solution: We have

$$\begin{aligned} \frac{d\mathbf{\Psi}_1^{-1}\mathbf{\Psi}_2}{dt} &= \frac{d\mathbf{\Psi}_1^{-1}}{dt}\mathbf{\Psi}_2 + \mathbf{\Psi}_1^{-1}\frac{d\mathbf{\Psi}_2}{dt} \\ &= -\mathbf{\Psi}_1^{-1}\frac{d\mathbf{\Psi}_1}{dt}\mathbf{\Psi}_1^{-1}\mathbf{\Psi}_2 + \mathbf{\Psi}_1^{-1}\mathbf{A}\mathbf{\Psi}_2 \\ &= -\mathbf{\Psi}_1^{-1}\mathbf{A}\mathbf{\Psi}_1\mathbf{\Psi}_1^{-1}\mathbf{\Psi}_2 + \mathbf{\Psi}_1^{-1}\mathbf{A}\mathbf{\Psi}_2 \\ &= -\mathbf{\Psi}_1^{-1}\mathbf{A}\mathbf{\Psi}_2 + \mathbf{\Psi}_1^{-1}\mathbf{A}\mathbf{\Psi}_2 \\ &= \mathbf{0} \end{aligned}$$

Therefore $\mathbf{\Psi}_1^{-1}\mathbf{\Psi}_2 = \mathbf{P}$ is a non-singular constant matrix and the result follows.

Exercise 6.7

1. (a) $\mathbf{x} = e^{5t}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

   (b) $\mathbf{x} = \frac{1}{2}e^{t}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - e^{-2t}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

   (c) $\mathbf{x} = e^{t}\begin{pmatrix} 1-3t \\ 4-3t \end{pmatrix}$

   (d) $\mathbf{x} = e^{t}\begin{pmatrix} 2t \\ 2t+1 \end{pmatrix}$

   (e) $\mathbf{x} = \frac{1}{4}e^{t}\begin{pmatrix} 6t-1 \\ 6t-3 \end{pmatrix} + \begin{pmatrix} t \\ 2t-1 \end{pmatrix}$

   (f) $\mathbf{x} = \frac{1}{2}\begin{pmatrix} -5\cos t - 5t\sin t \\ (t-2)\cos t - 2t\sin t \end{pmatrix}$

2. (a) $\mathbf{x}_p(t) = e^{5t}\mathbf{a}$

   (b) $\mathbf{x}_p(t) = e^{-2t}\mathbf{a} + e^{t}\mathbf{b}$

   (c) $\mathbf{x}_p(t) = e^{t}(t\mathbf{a} + \mathbf{b})$

   (d) $\mathbf{x}_p(t) = e^{t}(t\mathbf{a} + \mathbf{b})$

   (e) $\mathbf{x}_p(t) = e^{t}(t\mathbf{a} + \mathbf{b}) + t\mathbf{c} + \mathbf{d}$

   (f) $\mathbf{x}_p(t) = t(\cos t\mathbf{a} + \sin t\mathbf{b}) + \cos t\mathbf{c} + \sin t\mathbf{d}$